日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。
This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: 2003年 8月 6日

出願番号
Application Number: 特願2003-206168
ST. 10/C]: [JP2003-206168]

願人
Applicant(s): 株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

USSN 10/712,031
MATTINGLY, STANGER, MALUR & BRUNDIDG
(703) 684-1120
TSM-34

2003年10月15日

特許庁長官
Commissioner,
Japan Patent Office 今井康夫

出証番号 出証特2003-30845

BEST AVAILABLE COPY

【書類名】　特許願
【整理番号】　K03009841A
【あて先】　特許庁長官殿
【国際特許分類】　G06F 12/00
【発明者】
【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所システム開発研究所内
【氏名】　岩見　直子
【発明者】
【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所システム開発研究所内
【氏名】　山本　康友
【発明者】
【住所又は居所】　神奈川県川崎市麻生区王禅寺１０９９番地　株式会社日立製作所システム開発研究所内
【氏名】　藤本　和久
【特許出願人】
【識別番号】　000005108
【氏名又は名称】　株式会社　日立製作所
【代理人】
【識別番号】　100075096
【弁理士】
【氏名又は名称】　作田　康夫
【手数料の表示】
【予納台帳番号】　013088
【納付金額】　21,000円
【提出物件の目録】
【物件名】　明細書　1
【物件名】　図面　1

【物件名】　要約書　1

【プルーフの要否】　要

【書類名】　明細書

【発明の名称】　ストレージ装置

【特許請求の範囲】

【請求項１】

計算機と接続されるインターフェース部と、

ファイル操作の処理を行う第一の制御部と

記憶装置に対するデータの読み書きの処理を行う第二の制御部と、

前記インターフェース部、前記第一の制御部及び前記第二の制御部とを相互に接続する接続網とを有し、

前記インターフェース部は、前記計算機から送信されたフレームの転送先を前記第一の制御部及び前記第二の制御部のうちいずれか一つから選択し、前記接続網を介して前記フレームを選択した制御部へ転送することを特徴とするストレージ装置。

【請求項２】

前記第一の制御部は、前記インターフェース部から前記フレームを受信して前記フレームで指定されたファイル操作を行う場合、前記第二の制御部を介して前記記憶装置に対するデータの読み書きを実行することを特徴とする請求項１記載のストレージ装置。

【請求項３】

他のストレージ装置と接続する第二のインターフェース部を有し、

前記インターフェース部は、前記フレームの転送先を前記第一の制御部、前記第二の制御部又は前記第二のインターフェース部から選択することを特徴とする請求項２記載のストレージ装置。

【請求項４】

前記第一の制御部は、前記インターフェース部から前記フレームを受信して前記フレームで指定されたファイル操作を行う場合、前記第二のインターフェース部を介して前記他のストレージ装置に対するデータの読み書きを実行することを特徴とする請求項３記載のストレージ装置。

【請求項５】

前記第一の制御部を複数有し、

前記インターフェース部は、前記フレームがファイル操作を要求するコマンドを含むフレームであった場合、前記複数の第一の制御部から所定の第一の制御部を選択し、選択された前記第一の制御部へ前記フレームを前記接続網を介して送信することを特徴とする請求項１記載のストレージ装置。

【請求項６】

前記インターフェース部は、前記複数の第一の制御部と前記計算機から受信するフレームに含まれる識別子との対応関係に関する情報を保持し、前記フレームを受信した際に前記情報に基づいて前記フレームを転送する先の第一の制御部を決定することを特徴とする請求項５記載のストレージ装置。

【請求項７】

前記フレームを受信した前記第一の制御部の指示により、前記インターフェース部及び前記第二の制御部は、前記接続網を介して前記第一の制御部の処理に関するデータを送受信することを特徴とする請求項１記載のストレージ装置。

【請求項８】

前記フレームを受信した前記第一の制御部の指示により、前記インターフェース部及び前記第二のインターフェース部は、前記接続網を介して前記第一の制御部の処理に関するデータを送受信することを特徴とする請求項４記載のストレージ装置。

【請求項９】

前記情報には、前記フレームを受信する前記インターフェース部が有する一つのポートに複数の前記第一の制御部が対応することを示す情報が含まれていることを特徴とする請求項６記載のストレージ装置。

【請求項１０】

前記インターフェース部を複数有し、前記複数のインターフェース部で受信されたフレームが前記第一の制御部に転送されることを特徴とする請求項１記載のストレージ装置。

【請求項１１】

更に管理部を有し、前記インターフェース部は、前記管理部の指示に従って、

前記情報の内容を変更し、前記フレームの転送先を変更することを特徴とする請求項９記載のストレージ装置。

【請求項１２】

前記インターフェース部を複数有し、

前記管理部の指示に従って、前記インターフェース部で行っていた処理を他の前記インターフェース部に引継ぐ手段を有することを特徴とする請求項１１記載のストレージ装置。

【請求項１３】

前記管理部の指示は、前記インターフェース部の障害を該ストレージ装置が有する各装置が検出した際に行われ、前記管理部は、前記インターフェース部の障害の際に処理を引継ぐ前記他のインターフェース部の情報を有することを特徴とする請求項１２記載のストレージ装置。

【請求項１４】

前記第一の制御部を複数有し、

前記管理部の指示に従って、前記第一の制御部で行っていた処理を他の前記第一の制御部に引継ぐ手段を有することを特徴とする請求項１２記載のストレージ装置。

【請求項１５】

前記管理部の指示は、前記第一の制御部の障害を該ストレージ装置が有する各装置が検出した際に行われ、前記管理部は、前記第一の制御部の障害の際に処理を引継ぐ前記他の第一の制御部の情報を有することを特徴とする請求項１４記載のストレージ装置。

【請求項１６】

前記第二の制御部は、前記第二のインターフェース部を介して前記他のストレージ装置を制御することを特徴とする請求項４記載のストレージ装置。

【請求項１７】

前記第二の制御部は、キャッシュメモリ及びディスク装置を有することを特徴とする請求項１から１６に記載のうちいずれか一つのストレージ装置。

【請求項１８】

計算機と接続されるインターフェース部と、

ファイル操作の処理を行う制御部と

記憶装置に対するデータの読み書きの処理を行う第二のインターフェース部と

、

前記インターフェース部、前記第二のインターフェース部及び前記制御部とを相互に接続する接続網とを有し、

前記インターフェース部は、前記計算機から送信されたフレームの転送先を前記第一の制御部及び前記第二のインターフェース部のうちいずれか一つから選択し、前記接続網を介して前記フレームを選択した前記第一の制御部又は前記第二のインターフェース部へ転送することを特徴とするストレージ装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、計算機システムにおいて計算機によって使用されるデータを格納するストレージ装置、特にネットワーク経由でファイルアクセスされるストレージ装置に関する。

【０００２】

【従来の技術】

近年，計算機で取り扱われるデータ量は飛躍的に増大し，これに応じてデータを格納するストレージ装置の大容量化が進んでいる。大規模かつ低コストな大型ストレージを実現する方法として、特許文献１に開示されているクラスタ構成のストレージシステムがある。

【０００３】

クラスタ構成ストレージシステムは、比較的小構成のストレージ（以下「ストレージノード」）複数台をスイッチ等で相互に結合した構成を有する。このような構成では、ストレージノードを必要とされる記憶容量に合わせて複数結合してクラスタ構成ストレージシステムを構成することができ、小容量から大容量まで様々な容量を持つストレージシステムをスケーラブルに実現できる。

【０００４】

一方、ストレージ装置の種類の一つとして、ネットワークアタッチドストレージ（以下「NAS」）がある。これは、ネットワーク経由でのファイルアクセスを受け付けることが可能なストレージ装置で、実際には、ファイルシステムを提供する計算機（以下「NASサーバ」）及びディスク装置等の記憶装置から構成される。尚、「ファイル」とは、ファイルシステム等でユーザが認識するデータの一塊を指す。

NASは、使用者がファイルアクセスを直接利用できるため使いやすいストレージ装置である。しかし、NASへアクセスするためのプロトコル（ＮＦＳ等）が、TCP/IP(Transmission Control Protocol/Internet Protocol)の使用を前提としているため、NASへのアクセス頻度が上がるにつれ、NASにおけるTCP/IPの処理負荷が問題となってきた。この問題を解決する技術として、TCP/IPの処理をNASサーバから分離して単独のハードウエア（以下「TCP/IP処理部」）とし、TCP/IP処理部とNASサーバをスイッチで接続する技術が特許文献２に開示されている。特許文献２では、計算機が発行したコマンドは、TCP/IP処理部でTCP/IPレイヤまで処理され、その後NASサーバ部で処理される。NASサーバが記憶装置をアクセスする場合は、Fiber Channel Switching Module経由でSANに接続された記憶装置を利用する。

【０００５】

【特許文献１】

米国特許６，２５６，７４０号

【特許文献２】

ＰＣＴ国際公開　０２／０６９１６６号公報

【０００６】

【発明が解決しようとする課題】

特許文献２に開示されている技術では、アクセス対象となるNASサーバとNASサーバに付与されたネットワークアドレスが設定されるTCP/IP処理部のポートとの関係は１：１で一意に決定される。このため、NASサーバに対するアクセス頻度が上がったり、送受するデータ量が増加すると、TCP/IP処理部又はNASサーバの処理性能がネックとなり、レスポンスが低下する。また、TCP/IP処理部又はNAS

サーバに障害が発生した場合、障害が発生したNASの使用を継続できない。

又、上記の問題は、単純に特許文献１のクラスタ構成ストレージシステムにNASを導入するだけでは解決しない。

【０００７】

本発明の目的は、NASの使用負荷が上がった場合にも、レスポンスが低下しないNASを提供することにある。

また、本発明の別の目的は、NASに障害が発生した場合にも障害が発生したNASへのアクセスを継続できるNASを提供することにある。

【０００８】

【課題を解決するための手段】

本発明の一実施形態はストレージ装置であり、以下の構成を有する。ストレージ装置は、計算機との間で遣り取りされるフレーム（コマンド、データ等含む。パケット等でも良い）の通信プロトコルを処理するインターフェース部（以下「I/F処理部」）と、ファイル操作（ファイルへのデータの書き込み、読み出し、作成、削除等）の処理を行う第一の制御部（以下「NAS部」）と、記憶装置を制御する第二の制御部（以下「キャッシュ制御部」）と、各部を相互に接続する接続網とを有する。本構成において、計算機が送信したフレームを受信したI/F処理部は、受信したフレームをその内容に基づいてNAS処理部に送信したり、キャッシュ制御部に転送したりする。その際I/F処理部は、I/F処理部に割り当てられたアドレスとNAS部との対応関係の情報を用いる。

【０００９】

又、他の実施形態として、上述のストレージ装置が別のストレージ装置と接続される第二のI/F処理部を有する構成としても良い。尚、この場合、更に、上述の構成からキャッシュ制御部を省略した構成も考えられる。この場合、フレームは、第一のI/F処理部及び第二のI/F処理部を介して他のストレージ装置へ転送される。

【００１０】

一方、フレームを受け取ったNAS部は、状態に従い、キャッシュ制御部または、第二のI/F処理部経由で別のストレージ装置にアクセスして所定の処理を行っ

た後、フレームに対する応答をI/F処理部に送信する。その際NAS部は、ファイルと当該ファイルが記憶されているキャッシュ制御部又は他のストレージ装置の対応関係の情報を用いる。

【００１１】

なお、別の実施形態として、I/F処理部とNAS部を複数備える構成とし、I/F処理部とNAS部との対応関係を適宜設定することで、処理負荷の分散を図ったり、障害が発生した時に処理を引き継ぐ構成とすることが考えれる。

また、別の実施形態として、キャッシュ制御部がキャッシュメモリ及びディスク装置を有さない構成としても良い。

【００１２】

また、本実施例の好ましい実施形態として、I/F処理部に割り当てれらたアドレスとNAS部との対応関係、NAS部が処理するファイルと当該ファイルが記憶されているキャッシュ制御部もしは他のストレージ装置の対応関係等、各種情報を管理する管理用の計算機（以下、「管理サーバ」）又は同様の管理部をストレージ装置に設ける（又はストレージ装置に接続する）構成がある。管理サーバが、I/F処理部に割り当てれらたアドレスとNAS部との対応関係の変更命令をストレージ装置に出すことで、処理負荷の分散を図ったり、障害が発生した時に処理を引き継ぐよう構成変更をすることが可能な構成とする。

【００１３】

【発明の実施の形態】

図１は、本発明を適用したシステムの一実施形態の構成を示す図である。システムは、一台以上の計算機（以下「ホスト」）１００、一台以上のスイッチ１２０、ストレージ装置（以下「ストレージ」）１３０、ストレージ１３０とスイッチ１２０経由で接続される一台以上のストレージ（以下「外部ストレージ」）１８０、管理サーバ１１０及び管理端末１９０とを有する。

【００１４】

ホスト１００、スイッチ１２０、ストレージ１３０及び外部ストレージ１８０は、ＩＰネットワーク１７５を介して管理サーバ１１０に接続され、管理サーバ１１０で動作する図示しない管理ソフトウェアによって統合管理される。なお、

本実施形態では、ストレージ１３０は管理端末１９０を介して管理サーバ１１０に接続する形態をとるものとするが、ストレージ１３０が直接IPネットワークに接続される構成であっても良い。

【００１５】

ホスト１００は、ＣＰＵ１０１やメモリ１０２などを有する計算機であり、ディスク装置や光磁気ディスク装置などの記憶装置１０３に格納されたオペレーティングシステムやアプリケーションプログラムなどのソフトウェアをメモリ１０２に読み上げ、ＣＰＵ１０１がメモリ１０２から読み出して実行することで、所定の処理を行う。又、ホスト１００は、キーボードやマウスなどの入力装置１０４やディスプレイ等の出力装置１０５を具備し、入出力装置１０４でホスト管理者などからの入力を受け付け、出力装置１０５にCPUから指示された情報が出力される。

【００１６】

さらにホスト１００は、ストレージ１３０等外部の装置とネットワーク１７６を介してデータを送受するための一以上のポート１０７と、管理サーバとIPネットワーク１７５を介してデータ送受するための一以上のインタフェース制御部１０６を有する。

【００１７】

管理サーバ１１０は、ホスト１００と同様の計算機であり、システム全体の運用・保守管理といった所定の処理を行う。CPU１１１によって管理ソフトウエアが実行されると、管理サーバ１１０はインタフェース制御部１１６からＩＰネットワーク１７５を介して、システム内の各機器から構成情報、リソース利用率、性能監視情報などを収集する。そして、管理サーバ１１０は収集したそれらの情報を出力装置１１５に出力してストレージの管理者等に提示する。また、管理サーバは、入力装置１１４を介してストレージ管理者からの指示を受信し、受信した運用・保守指示をインタフェース制御部１１６を介して各機器に送信する。

【００１８】

スイッチ１２０は、複数のポート１２１及びインターフェース制御部を有する。ポート１２１は、ホスト１００のポート１０７とIPネットワーク１７６を介し

て、または、ストレージ１３０のポート１４１とIPネットワーク１７７を介して接続される。又、インタフェース制御部１２３は、ＩＰネットワーク１７５を介して管理サーバ１１０に接続される。スイッチ１２０は、IPネットワーク内に存在するスイッチが有する経路制御等一般的な処理を行う。

【００１９】

ストレージ１３０はクラスタ構成を有する。具体的には、ストレージ１３０は複数のプロトコル変換部１４０、複数のキャッシュ制御部１５０、１つ以上のNAS部１４５及び構成管理部１６０を有し、これらは相互結合網１７０で相互に接続されている。

【００２０】

プロトコル変換部１４０は、複数のポート１４１、１つ以上の制御プロセッサ１４２、メモリ１４３、及び相互結合網１７０と接続される転送制御部１４４を有する。プロトコル変換部１４０は、ポート１４１から受信したコマンドを解析し、アクセス対象となるデバイスやNAS部１４５を特定し、転送制御部１４４より相互結合網１７０を介して適当なキャッシュ変換部１５０やNAS部１４５等へコマンドやデータを転送する。尚、以下「デバイス」とは、ストレージが提供する物理的又は論理的な記憶領域のことを言う。

【００２１】

また、プロトコル変換部１４０は、ポート１４１を介して外部ストレージ１８０など別のストレージと接続されることが可能である。この場合、外部ストレージ１８０と接続されるプロトコル変換部１４０は、別のプロトコル変換部１４０から受信する入出力要求やキャッシュ制御部１５０から受信する入出力要求を外部ストレージ１８０へ送信する。これにより、ホスト１００又はNAS部145は、外部ストレージ１８０にデータを書き込んだり（以下「ライト」）、外部ストレージ１８０からデータを読み出したり（以下「リード」）することができる。

【００２２】

なお本実施形態では、ポート１４１はIPネットワークで主に使用されるTCP/IPと、SCSIを上位プロトコルとしたFiber Channel（以下「FC」）に対応しているとする。NASへのアクセスに使用されるNFS等の通信プロトコルは、TCP/IPプロト

コルの上位プロトコルとして位置する。本実施形態では、記述の簡単のため、NASへのアクセスに使用されるファイルアクセスプロトコルをNFSとするが、CIFS等別のプロトコルでもかまわないし、複数のファイルアクセスプロトコルをサポートしてもかまわない。また、IPネットワークを介してストレージアクセス（デバイスのアクセス）を行う場合に使用されるiSCSI(Internet SCSI)等も、TCP/IPプロトコルの上位プロトコルとして位置し、プロトコル変換部１４０は、NFS等NASのアクセスに使用されるプロトコルに加え、iSCSI等ストレージが有するデバイスをアクセスするためのプロトコルに対応する。

【００２３】

ホスト１００がストレージ１３０又は外部ストレージ１８０のデバイスをブロックデバイスとして利用したい場合、FC又はIPのネットワーク経由でFCやiSCSI等のプロトコルを用いてストレージ１３０と通信を行う。

なお、ポート１４１がFCのプロトコルに対応しない構成でも良い。また、NFS等NASのアクセスに使用されるプロトコルのみに対応する構成でも良い。

【００２４】

キャッシュ制御部１５０は、1つ以上のポート１５６、各々ポートに接続される１台以上のディスク装置１５７、１つ以上の制御プロセッサ１５２、各々の制御プロセッサ１５２に対応するメモリ１５３、1つ以上のディスクキャッシュ１５４、1つ以上の制御メモリ１５５及び相互結合網１７０と接続される転送制御部１５１を有する。

【００２５】

制御プロセッサ１５２は、相互結合網１７０を介して転送制御部１５１で受信した、同じキャッシュ制御部１５０内に存在するディスク装置１５７への入出力要求を処理する。また制御プロセッサ１５２は、キャッシュ制御部１５０がホスト１００またはNAS部１４５に対してディスクアレイのように複数のディスク装置１５７を１つ又は複数の下位論理デバイスとして提供している場合に、下位論理デバイスとディスク装置１５７との間の対応関係の管理や、下位論理デバイスに対するアクセス要求をディスク装置１５７へのアクセス要求に変換する処理などを行う。更に制御プロセッサ１５２はデータ複製やデータ再配置などの各種処

理を実行する。

【００２６】

ディスクキャッシュ１５４には、ホスト１００またはNAS部１４５からのアクセス要求に対する処理速度を高めるため、ディスク装置１５７から頻繁に読み出されるデータが予め格納されたり、ホスト１００から受信したライトデータが一時的に格納される。尚、ディスクキャッシュ１５４を用いたライトアフタ、即ちホスト１００またはNAS部１４５から受信したライト用のデータ（以下「ライトデータ」）がディスクキャッシュ１５４に格納された後、ディスク装置１５７に実際に書き込まれる前にホスト１００またはNAS部１４５に対しライト要求に対する応答を返す場合、ディスクキャッシュ１５４に格納されているライトデータがディスク装置１５７に書き込まれる前に消失することを防止するため、ディスクキャッシュ１５４をバッテリバックアップなどにより不揮発メモリとしたり、媒体障害への耐性向上のため二重化するなど、ディスクキャッシュ１５４の可用性を向上させておくことが望ましい。

【００２７】

制御メモリ１５５には、キャッシュ制御部１５０が、ディスク装置１５７、複数のディスク装置１５７を組み合わせて構成される物理デバイス又はストレージ１３０にプロトコル変換部１４０を介して接続される外部ストレージ１８０内のデバイス（以下「外部デバイス」）の管理や、外部デバイス又は物理デバイスと下位論理デバイスの対応関係の管理を行うために必要な情報が格納される。なお、制御メモリ１５５に格納されている制御情報が消失すると、ホスト１００やNAS部１４５はキャッシュ制御部１５０が有するディスク装置１５７に格納されているデータへアクセスできなくなるため、制御メモリ１５５はバッテリバックアップなどにより不揮発メモリとしたり、媒体障害への耐性向上のため二重化するなど、高可用化のための構成を有することが望ましい。

【００２８】

また、キャッシュ制御部１５０の別の構成として、ポート１５６とディスク１５７を有さず、外部ストレージ１８０のデバイスのみを制御する構成としても良い。

【００２９】

NAS部１４５は、一または複数のCPU１４７、各々の制御プロセッサに対応する一または複数のメモリ１４９、記憶装置１４８及び相互結合網１７０と接続される転送制御部１４６を有する。記憶装置１４８に格納された制御プログラムをメモリ１４９に読み上げ、これをCPU１４７が実行することで、NAS部１４５は所定の処理を行う。

【００３０】

CPU１４７は、相互結合網１７０を介して転送制御部１４６より受信したコマンドを処理し、コマンドを送信してきたプロトコル変換部１４０へ応答コマンドを送信する。また、CPU１４７は、コマンド処理時にデバイスアクセスが必要になった場合（ファイルへのデータの書き込みや読み出し）、そのデバイスを提供（より具体的には、NAS部で実行されているファイルシステムにマウントされるデバイスを提供する）しているキャッシュ制御部１５０又は外部ストレージ１８０に対し入出力要求を転送制御部１４６を介して送信し、その応答を相互結合網１７０を介して転送制御部１４６より受信する。尚、デバイスへアクセスを行う場合、NAS部145は、受信したファイルアクセスのコマンドをそのファイルのデータが格納されるデバイスのブロックへのアクセスコマンドに変換する。

【００３１】

構成管理部１６０は、一又は複数の制御プロセッサ１６２、各々の制御プロセッサに対応する一又は複数のメモリ１６３、一又は複数の制御メモリ１６４、記憶装置１６５、相互結合網１７０と接続される転送制御部１６１及びインタフェース制御部１６６を有する。記憶装置１６５に格納された制御プログラムをメモリ１６３に読み上げ、これを制御プロセッサ１６２が実行することで、ストレージ１３０の構成管理や障害管理のための所定の処理が実行される。

【００３２】

制御プロセッサ１６２は、インタフェース制御部１６６を介して接続される管理端末１９０にストレージ管理者へ提示する構成情報を送信したり、管理者から管理端末１９０に入力された保守・運用指示を管理端末１９０から受信して、受領した指示に従い、ストレージ１３０の構成変更などを行う。ストレージ１３０

の構成情報は制御メモリ１６４に格納される。制御メモリ１６４に格納された構成情報はプロトコル変換部１４０の制御プロセッサ１４２、キャッシュ制御部１５０の制御プロセッサ１５２、NAS部１４５の制御プロセッサ１４７によって参照されたり更新されたりすることができる。これにより、ストレージ１３０内のプロトコル変換部１４０及びキャッシュ制御部１５０間で構成情報を共有することができる。

【００３３】

なお、構成管理部１６０が障害などにより動作不可に陥った場合、ストレージ１３０全体へのアクセスが不可能になるため、構成管理部１６０内の各構成要素を二重化しておくか、もしくは構成管理部１６０自体をストレージ１３０内に複数搭載して構成管理部１６０自体の二重化をしておくことが望ましい。また、管理端末１９０から個々のキャッシュ制御部１５０へのＩ／Ｆを別途設け、構成管理部１６０が行う制御を個々のキャッシュ制御部１５０と管理端末１９０で分担することで、構成管理部１６０を制御メモリ１６４のみの構成にすることも可能である。さらに、制御メモリ１６４内の情報を個々のキャッシュ制御部１５０の制御メモリ１５５で保持することで、構成管理部１６０を省略することも可能である。

【００３４】

相互結合網１７０は、プロトコル変換部１４０、キャッシュ制御部１５０及び構成管理部１６０を相互に接続し、これらの各装置間でのデータ、制御情報及び構成情報の送受信を実行する。この相互結合網１７０により、構成管理部１６０がプロトコル変換部１４０、キャッシュ制御部１５０およびNAS部１４５にストレージ１３０の構成情報を配布したり、プロトコル変換部１４０、キャッシュ制御部１５０及びNAS部１４５から構成情報を取得してストレージ１３０の構成を管理したりすることができる。

【００３５】

また、相互結合網１７０がプロトコル変換部１４０、キャッシュ制御部１５０及びNAS部１４５との間のアクセス要求を転送するので、ホスト１００はプロトコル変換部１４０の任意のポート１４１からNAS部１４５やキャッシュ制御部１

５０に属する任意の上位論理デバイスへアクセスすることが可能となる。なお、可用性向上の観点から相互結合網１７０も多重化されていることが望ましい。

【００３６】

管理端末１９０は、CPU１９２、メモリ１９３、記憶装置１９４、構成管理部１６０と接続するインタフェース制御部１９１、IPネットワーク１７５と接続するインタフェース制御部１９７、ストレージ管理者からの入力を受け付ける入力装置１９５及びストレージ管理者にストレージ１３０の構成情報や管理情報を出力するディスプレイ等の出力装置１９６を有する計算機である。CPU１９２は記憶装置１９４に格納されているストレージ管理プログラムをメモリ１９３に読み出して、これを実行することにより、構成情報の参照、構成変更の指示、特定機能の動作指示などを行い、ストレージ１３０の保守運用に関して、ストレージ管理者もしくは管理サーバ１１０とストレージ１３０間のインタフェースとなる。

【００３７】

なお、管理端末１９０を省略して、ストレージ１３０を直接管理サーバ１１０へ接続し、ストレージ１３０を管理サーバ１１０で動作する管理ソフトウェアを用いて管理してもよい。

【００３８】

外部ストレージ１８０は、一又は複数のポート１８１、制御プロセッサ１８２、メモリ１８３、ディスクキャッシュ１８４、一又は複数のディスク装置１８６及び各々ディスク装置１８６に接続される一又は複数のポート１８５を有する。外部ストレージ１８０は、各々が有するポート１８１及びスイッチ１２０を介してストレージ１３０に接続される。又、外部ストレージ１８０が有するディスク装置１８６から構成される論理的な記憶領域である外部デバイスは、ストレージ１３０によりストレージ１３０が有する上位論理デバイスとしてNAS部１４５またはホスト１００へ提供される。

【００３９】

なお、システムはスイッチ１２０の他にFCスイッチを有し、FCスイッチ経由で外部ストレージ１８０がストレージ１３０と接続されてもよい。また、スイッチを経由せず直接ストレージ１３０と外部ストレージ１８０とを接続してもよい。

また、外部ストレージ１８０を設けず使用しないシステム構成としてもよい。

【００４０】

制御プロセッサ１８２はメモリ１８３に格納されているプログラムを実行することにより、ポート１８１から受信したディスク装置１８６への入出力要求を処理する。本実施形態では外部ストレージ１８０を、クラスタ構成を有さず、ストレージ１３０より小規模な構成のストレージ装置とするが、ストレージ１３０と同じ構成を有する同規模のストレージ装置であってもかまわない。また、外部ストレージ１８０がディスクキャッシュ１８４や制御プロセッサ１８２の無いＪＢＯＤ（Ｊｕｓｔ　Ｂｏｕｎｃｈ　Ｏｆ　Ｄｉｓｋ）のような単体ディスク装置グループであっても構わない。

【００４１】

以下、本実施形態のシステムの動作概要について簡単に説明する。

ストレージ１３０の有するディスク装置１５７または外部デバイスからなる上位論理デバイスの一部もしくは全ては、ファイルアクセス処理を行うNAS部１４５に提供される。

【００４２】

ホスト１００がストレージ１３０に記憶されるファイルを使用するためにストレージ１３０に送信したコマンド（ファイルアクセスのコマンド）は、ストレージ１３０のポート１４１の内、ファイルアクセス用（NFS）のアドレスが割り当てれらているポート１４１に到着する。そのポート１４１を有するプロトコル変換部１４０は、そのコマンドを解析し、そのコマンドを処理すべきNAS部１４５を選択し、そのコマンドを相互結合網１７０経由で選択したNAS部１４５に送信する。

【００４３】

コマンドを受信したNAS部１４５は、そのコマンドで指定されるファイルに対する処理を行う。コマンドの処理中にデバイス（具体的にはそのデバイスに格納されているファイル）へのアクセスが必要となった場合、そのデバイスがキャッシュ制御部１５０が提供する上位論理デバイスであれば、NAS部１４５は当該キャッシュ制御部１５０へアクセスコマンドを送信し、必要なデータを得たり、デ

ータをキャッシュ制御部１５０に格納したりする。尚、上位論理デバイスに外部デバイスが含まれる場合、キャッシュ制御部１４５は、プロトコル変換部１４５を介してコマンドを外部ストレージ１８０へ送信する。また、NAS部１４５が外部ストレージ１８０が提供するデバイスを直接利用している場合、NAS部１４５は、当該外部ストレージが接続されているプロトコル変換部１４０にアクセスコマンドを送信し、必要なデータを得たり、データを外部ストレージ１８０が提供するデバイスに格納する。

【００４４】

コマンドの処理を終了したNAS部１４５は、応答コマンドを作成し、相互結合網１７０を経由して先に受信したコマンドの送信元のプロトコル変換部１４０に応答コマンドを送信する。応答コマンドを受信したプロトコル変換部１４０は、応答コマンドをホスト１００へ送信する。

【００４５】

なお、以降の説明では、表現の簡略化のため、プロトコル変換部１４０をＰＡ、キャッシュ制御部１５０をＣＡ、構成管理部１６０をＭＡ、管理端末１９０をＳＴ、NAS部１４５をＡＡと表記する。

【００４６】

また、本発明の実施形態では、ストレージ１３０内の各ディスク装置、物理デバイス、下位又は上位論理デバイスの関係（以下「デバイス階層」）を次の通りとする。

まず、ＣＡ内でディスク装置１５７複数台によるディスクアレイが構成され、このディスクアレイはＣＡによって物理デバイスとして管理される。また、ＰＡに接続された外部ストレージ１８０が提供する外部デバイスは、ＰＡで認識された後、ＭＡにて管理される。

【００４７】

更にＣＡでは、ＣＡ内に存在する物理デバイス又は当該ＣＡを介してアクセスされる外部デバイスに対して下位論理デバイスを割り当てる。下位論理デバイスは各ＣＡ内において管理される論理デバイスであり、その番号は各ＣＡ毎に独立に管理される。下位論理デバイスは、ＭＡによって上位論理デバイスと対応付け

られ、ストレージ１３０が提供する論理的な記憶装置（デバイス）としてホスト１００またはＡＡに提供される。

【００４８】

即ち、ホスト１００またはＡＡが認識するのはストレージ１３０の上位論理デバイスであり、ホスト１００は上位論理デバイスを識別するための情報を用いて、ストレージ１３０若しくは外部ストレージ１８０に格納されているデータにアクセスする。ただし、ＡＡは、管理情報に従いＰＡを介して直接外部ストレージ１８０の外部デバイスにアクセスすることもできる。この場合、外部デバイスは特定のＣＡには割り当てられていない。

【００４９】

尚、個々のディスク装置１５７を１つの物理デバイスおよび１つの下位論理デバイス及び上位論理デバイスとしてもよい。また、複数の物理デバイスを1つ又は複数の下位論理デバイスに対応付けたり、複数の下位論理デバイスに１つ又は複数の上位論理デバイスを対応付けても良い。

以下、本システムの動作を説明する前に、本システムの各構成要素が有する情報について説明する。

【００５０】

図２は、ストレージ１３０の各構成要素の制御メモリおよびメモリに格納される、制御情報、管理情報とストレージ制御処理のためのプログラムの一例を示したソフトウェア構成図である。

【００５１】

ストレージ１３０のデバイス階層を管理するための管理情報として、ＣＡの制御メモリ１５５に下位論理デバイス管理情報２０１、物理デバイス管理情報２０２及びキャッシュ管理情報２０３が、ＭＡの制御メモリ１６４に上位論理デバイス管理情報２０４、外部デバイス管理情報２０５及びＬＵパス管理情報２０６が格納されている。

【００５２】

ＣＡ内の制御メモリ１５５およびＭＡに格納されている各制御情報は、各ＣＡ、各ＰＡ、各ＡＡ、及びＭＡ内の各制御プロセッサから参照・更新可能であるが

、その際に相互接続網１７０などを介したアクセスが必要となる。よって、処理性能向上のため、各制御プロセッサで実行される処理に必要な制御情報の複製が、各部位（即ち、CA、PA、AA及びMA）内のメモリに保持される。

【００５３】

各部位は、当該部位が管理する制御情報が構成変更により更新された場合は、相互接続網１７０を介してその旨を他の部位に通知し、最新情報を当該部位の制御メモリから他の部位のメモリへ取り込ませる。なお、制御情報が更新された場合に更新を他の部位に知らせる別の方法として、例えば、各部位は、自部位内の制御メモリに保持した構成情報毎に制御メモリ上に更新有無を示すフラグを設け、各部位の制御プロセッサは処理開始時もしくは各構成情報を参照する毎に同フラグを参照して更新有無をチェックする方法などが考えられる。

【００５４】

なお、各部位のメモリには上述の制御情報の複製に加えて、各部位内の制御プロセッサで動作する制御プログラムが格納される。

【００５５】

NASの処理（ファイルアクセス処理、障害時のフェイルオーバ等）にかかわる管理情報としては、ＭＡの制御メモリ１６４にNAS管理情報２２３、ポート管理情報２２４が格納されている。

【００５６】

ＰＡのメモリ１４３には、上位論理デバイス管理情報２０４の全て又は当該ＰＡの処理にかかわる情報の複製２１４、外部デバイス管理情報２０５の全て又は当該ＰＡの処理にかかわる情報の複製２１５、LUパス管理情報２０６の全て又は当該ＰＡの処理にかかわる情報の複製２１６、ポート管理情報２２４の全て又は当該ＰＡの処理にかかわるポート管理情報の複製２２２、NAS管理情報２２３の全て又は当該ＰＡの処理にかかわるNAS管理情報の複製２２１、ポート１４１または転送制御部１４４から受信したデータもしくはコマンドを振り分ける要求振り分け処理プログラム２５１、LUパス定義処理プログラム２５２、外部デバイス定義処理プログラム２５３、NAS／ポート障害処理プログラム２３１及びファイルアクセス処理プログラム２２９が格納されている。

【００５７】

ＣＡの制御メモリ１５５には、下位論理デバイス管理情報２０１、キャッシュ管理情報２０３、物理デバイス管理情報２０２が格納される。メモリ１５３には、下位論理デバイス管理情報２０１の複製２１１、物理デバイス管理情報２０２の複製２１２、外部デバイス管理情報２０５の全て又は当該ＣＡの処理にかかわる情報の複製２１５、ＰＡまたはＡＡから受信した入出力コマンドを処理するコマンド処理プログラム２５４、論理デバイス定義処理プログラム２５５、ライトアフタ処理プログラム及び外部デバイス接続変更処理プログラム２５６が格納されている。

【００５８】

ＭＡの制御メモリ１６４には、上位論理デバイス管理情報２０４、外部デバイス管理情報２０５、LUパス管理情報２０６、NAS管理情報２２３及びポート管理情報２２４が格納される。メモリ１６３には、全デバイス管理情報(即ち、上位論理デバイス管理情報２０４、下位論理デバイス管理情報２０１、物理デバイス管理情報２０２及び外部デバイス管理情報２０５)の複製２１０、論理デバイス定義処理プログラム２５５、外部デバイス接続変更処理２５６、LUパス定義処理２５２、NAS/ポート障害処理プログラム２３１、外部デバイス定義処理プログラム２５３、NAS管理情報変更処理プログラム２２５及びポート管理情報変更処理プログラム２２６が格納されている。

【００５９】

ＡＡのメモリ１４９には、NAS管理情報２２３の全て又は当該ＡＡの処理にかかわる情報の複製２２７、外部デバイス管理情報２０５の全て又は当該ＡＡの処理にかかわる情報の複製２１５、当該ＡＡのキャッシュ管理情報、NAS/ポート障害処理プログラム２３１及びファイルアクセス処理プログラム２２９が格納されている。

【００６０】

なお、本実施形態では、全ての管理情報（下位論理デバイス管理情報２０１、物理デバイス管理情報２０２、上位論理デバイス管理情報２０４、外部デバイス管理情報２０５、LUパス管理情報２０６、NAS管理情報２２３、ポート管理情報

２２４及び各複製）は、ストレージ１３０の工場出荷時にあらかじめ、初期値が設定されているものとする。その後、適宜、ストレージ管理者または、管理サーバからの制御により後述の方法を用いて値が設定変更され、運用される。

【００６１】

図１５は、ＳＴの制御メモリに格納される、制御情報とストレージ制御処理のためのプログラムの一例を示したソフトウェア構成図である。メモリ１９３には、全デバイス管理情報(即ち、上位論理デバイス管理情報２０４、下位論理デバイス管理情報２０１、物理デバイス管理情報２０２及び外部デバイス管理情報２０５)の複製２１０、NAS管理情報２２３の複製２３１、ポート管理情報２２４の複製２３２、論理デバイス定義処理プログラム２５５、LUパス定義処理プログラム２５２、外部デバイス定義処理プログラム２５３、NAS管理情報変更処理プログラム２２５、ポート管理情報変更処理プログラム２２６及びNAS／ポート障害処理プログラム２３１が格納される。

【００６２】

図３は、上位論理デバイス管理情報２０３の一構成例を示す図である。上位論理デバイス管理情報は、各上位論理デバイス毎に、上位論理デバイス番号から対応外部デバイス番号までの情報の組を保持するエントリ３１～３９を有する。なお、切替進捗ポインタを使用しない場合はエントリ３９を有しなくても良い。

【００６３】

エントリ３１には、対応する上位論理デバイスを識別するために上位論理デバイスにＭＡが割り当てた番号が格納される。エントリ３２には、上位論理デバイス番号により特定される上位論理デバイスの記憶容量が格納される。エントリ３３には、当該上位論理デバイスに対応付けられている下位論理デバイスの番号と当該下位論理デバイスが属するＣＡの番号が格納される。上位論理デバイスが未定義の場合、エントリ３３には無効値が設定される。尚、この下位論理デバイス番号が、当該下位論理デバイスを管理するＣＡが保持する下位論理デバイス管理情報２０１のエントリ番号となる。

【００６４】

エントリ３４には、対応する上位論理デバイスの状態を示す情報が設定される

。状態としては、「オンライン」、「オフライン」、「未実装」及び「障害オフライン」が存在する。「オンライン」は、当該上位論理デバイスが正常に稼動し、ホスト１００またはＡＡからアクセスできる状態であることを示す。「オフライン」は、当該上位論理デバイスは定義され、正常に稼動しているが、ＬＵパスが未定義であるなどの理由で、ホスト１００またはＡＡからはアクセスできない状態にあることを示す。「未実装」は、当該上位論理デバイスが定義されておらず、ホスト１００またはＡＡからアクセスできない状態にあることを示す。「障害オフライン」は、当該上位論理デバイスに障害が発生してホスト１００またはＡＡからアクセスできないことを示す。

【００６５】

なお、本実施形態では簡単のため、ストレージ１３０の工場出荷時にあらかじめ、ディスク装置１５７で作成された物理デバイスへ下位論理デバイスが割り当てられ、更に下位論理デバイスに上位論理デバイスが割り当てられているものとする。このため、下位論理デバイスに割当済みの上位論理デバイスについてのエントリ３４の初期値は「オフライン」、その他の上位論理デバイスについては上位論理デバイスが定義されていないことになるので、エントリ３４の初期値は「未実装」となる。

【００６６】

エントリ３５には、ポート番号、ターゲットＩＤ及びＬＵＮが登録される。ここで登録されるポート番号は、対応する上位論理デバイスが複数のポート１４１またはＡＡの転送制御部１４６のうちどのポートまたはＡＡの転送制御部１４６に接続されているかを表す情報、即ち対応する上位論理デバイスにアクセスするために用いられるポート１４１またはＡＡの転送制御部１４６の識別情報である。ここでポート１４１又はＡＡの転送制御部１４６の識別情報は、各ポート１４１及びＡＡの転送制御部１４６に割り振られているストレージ１３０内で一意な番号であり、エントリ３５には、当該上位論理デバイスがＬＵＮ定義されているポート１４１又はＡＡの転送制御部１４６の番号が記録される。

【００６７】

また、ターゲットＩＤとＬＵＮは、当該上位論理デバイスを識別するための識

別子である。本実施形態においては、上位論理デバイスを識別するための識別子として、ＳＣＳＩプロトコルでホスト１００またはＡＡからデバイスをアクセスする場合に用いられるＳＣＳＩ－ＩＤ及び論理ユニット番号（ＬＵＮ）を用いる。

【００６８】

エントリ３６には、当該上位論理デバイスへのアクセスが許可されているホスト１００またはＡＡを識別するホスト名が登録される。ホスト名としては、ホスト１００のポート１０７に付与されたWWN(World Wide Name)やiSCSIネームなど、ホスト１００又はポート１０７及びＡＡを一意に識別可能な値であれば何を用いてもよい。ストレージ１３０は、このほかに、各ポート１４１のWWNやMAC (Media Access Control)アドレス、IPアドレスなどの属性に関する管理情報を保持する。

【００６９】

エントリ３７には、当該上位論理デバイスへの入出力要求の処理形態を示す「ＣＡ経由」、「ＰＡ直結」又は「ＰＡ直結移行中」という値が格納される。

「ＣＡ経由」は、当該上位論理デバイスが外部デバイスと対応付けられていない場合や、当該上位論理デバイスが外部デバイスに対応付けられていて、かつ当該上位論理デバイスに対する入出力要求に対してＣＡ１５０での入出力処理が必要な処理形態を示す。

【００７０】

「ＰＡ直結」は、当該上位論理デバイスが外部デバイスに対応付けられており、かつ、ＣＡを経由せずに直接ＰＡ間で当該上位論理デバイスへの入出力要求を転送できる処理形態を示す。

「ＰＡ直結移行中」は、外部デバイス接続形態を「ＣＡ経由」から「ＰＡ直結」へ切り替える過渡状態、すなわち、後述する切替処理が対応ＣＡにて実行中であることを示す。

【００７１】

エントリ３８には、当該上位論理デバイスが外部デバイスに対応付けられている場合にはその外部デバイスの識別番号が、そうでない場合は無効値が設定され

る。

【００７２】

エントリ３９には、切替進捗ポインタが登録される。切替進捗ポインタは、当該上位論理デバイスが「ＰＡ直結移行中」である場合に使用され、当該上位論理デバイスが有する記憶領域のうち、外部接続の接続形態変更処理が未完了な領域の先頭アドレスを示す情報である。なお、「ＰＡ直結移行中」処理を行わない場合には、エントリ３９を省いてもかまわない。

【００７３】

図４は、ＬＵパス管理情報２０６の一構成例を示す図である。LUパス管理情報２０６は、ＰＡが有する各ポート１４１又はＡＡの転送制御部１４６に対して定義されている有効なＬＵＮ毎に、情報を保持する為のエントリ４１～４３を有する。エントリ４１には、ポート１４１又はＡＡの転送制御部１４６に定義された(割り当てられた)ＬＵＮのアドレスが格納される。

【００７４】

エントリ４２には、当該ＬＵＮが割り当てられている上位論理デバイスの番号が格納される。エントリ４３には、当該ＬＵＮに対してアクセスを許可されているホスト１００またはＡＡを示す情報、具体的にはWWNや転送制御部１４６の番号等が登録される。

【００７５】

尚、一つの上位論理デバイスに対して複数のポート１４１またはＡＡの転送制御部１４６が定義され（割り当てられ）ており、複数のポート１４１またはＡＡから当該上位論理デバイスにアクセスできる場合がある。この場合、当該複数のポート１４１またはＡＡのＬＵＮ各々に関するLUパス管理情報２０６のエントリ４３に登録された値の和集合が、当該上位論理デバイスに関する上位論理デバイス管理情報２０３のエントリ３６に保持される。

【００７６】

図５は、下位論理デバイス管理情報２０１の一構成例を示す図である。各ＣＡは、自ＣＡに属する各下位論理デバイス毎に、下位論理デバイス番号から対応上位論理デバイス番号までの情報の組を登録するエントリ５１～５６を有する。な

お、切替進捗ポインタが使用されない場合は、エントリ５６は省いてもかまわない。

【００７７】

エントリ５１には、対応する下位論理デバイスを識別するための識別番号が登録される。エントリ５２には、エントリ５１により特定される下位論理デバイスの記憶容量が格納される。

【００７８】

エントリ５３には、当該下位論理デバイスに対応付けられる、ＣＡの物理デバイスの識別番号又は外部ストレージ１８０に存在する外部デバイスの識別番号が格納される。当該下位論理デバイスに対して物理デバイス又は外部デバイスが割り当てられていない場合、エントリ５３には無効値が設定される。なお、エントリ５３に登録されるデバイス番号は、当該下位論理デバイスを管理しているＣＡが保持する物理デバイス管理情報２０２、もしくはＭＡが保持する外部デバイス管理情報２０５のエントリ番号と一致する。

【００７９】

エントリ５４には、当該下位論理デバイスの状態を示す情報が設定される。設定される情報の内容は、上位論理デバイス管理情報２０３のエントリ３４と同じであるため、説明は省略する。エントリ５５には、当該下位論理デバイスに対応付けられている上位論理デバイス番号が設定される。

【００８０】

エントリ５６には、切替進捗ポインタが登録される。このポインタは上位論理デバイス管理情報２０４の切替進捗ポインタと同じで、当該下位論理デバイスに対応する上位論理デバイスが「ＰＡ直結移行中」である場合に使用され、移行中の上位論理デバイスに対応する下位論理デバイスのうち、外部接続の接続形態変更処理が未完了な記憶領域の先頭アドレスを示す情報である。

【００８１】

図６は、ＣＡのディスク装置１５７から構成される物理デバイスを管理するための物理デバイス管理情報２０２の一構成例を示す図である。各ＣＡは、自ＣＡが有する物理デバイス毎に、物理デバイス番号からディスク内サイズの情報の組

を保持する為のエントリ６１～６９を有する。

【００８２】

エントリ６１には、物理デバイスを識別するための識別番号が登録される。エントリ６２には、エントリ６１に登録された値で特定される物理デバイスの記憶容量が格納される。エントリ６３には、当該物理デバイスに対応付けられる下位論理デバイスの番号が格納される。当該物理デバイスが下位論理デバイスへ割り当てられていない場合、エントリ６３には無効値が設定される。

【００８３】

エントリ６４には、当該物理デバイスの状態を示す情報が設定される。状態としては、「オンライン」、「オフライン」、「未実装」及び「障害オフライン」が存在する。「オンライン」は、当該物理デバイスが正常に稼動し、下位論理デバイスに割り当てられている状態であることを示す。「オフライン」は、当該物理デバイスは定義され正常に稼動しているが、下位論理デバイスに未割り当てであることを示す。「未実装」は、当該物理デバイスがディスク装置１５７に対応付けされていない状態にあることを示す。「障害オフライン」は、当該物理デバイスに障害が発生して下位論理デバイスに割り当てられないことを示す。

【００８４】

なお、本実施形態では、簡単のため、物理デバイスは製品の工場出荷時にあらかじめディスク装置１５７に対応付けられて作成され、かつ下位論理デバイスにも割り当てられているものとする。このため、利用可能な物理デバイスについてはエントリ６４の初期値は「オンライン」状態、その他は「未実装」状態となる。

【００８５】

エントリ６５には、当該物理デバイスが割り当てられたディスク装置１５７のＲＡＩＤレベル、データディスクとパリティディスク数などＲＡＩＤ構成に関連する情報が保持される。同じように、エントリ６６には、ＲＡＩＤにおけるデータ分割単位(ストライプ)長が保持される。エントリ６７には、当該物理デバイスが割り当てられたＲＡＩＤを構成する複数のディスク装置１５７各々の識別番号が保持される。このディスク装置１５７の識別番号は、ＣＡ内でディスク装置１

５７を識別するために付与された一意な値である。

【００８６】

エントリ６８とエントリ６９には、当該物理デバイスが各ディスク装置１５７内のどの領域に割り当てられているかを示す情報（開始オフセット及びサイズ）が登録される。本実施例では簡単のため、全物理デバイスについて、ＲＡＩＤを構成する各ディスク装置１５７内のオフセットとサイズを統一している。

【００８７】

図７は、ストレージ１３０に接続された外部ストレージ１８０の外部デバイスを管理するための、外部デバイス管理情報２０５の一構成例を示す図である。ストレージ１３０のMAは、外部ストレージ１８０が有する外部デバイス毎に、外部デバイス番号からターゲットポートＩＤ／ターゲットＩＤ／ＬＵＮリストまでの情報の組を保持する為のエントリ７１～７９を有する。

【００８８】

エントリ７１には、ストレージ１３０のＭＡが当該外部デバイスに対して割り当てた、ストレージ１３０内で一意な値が格納される。エントリ７２には、エントリ７１に登録された値より特定される外部デバイスの記憶容量が格納される。

【００８９】

エントリ７３には、当該外部デバイスに対応付けられているストレージ１３０内の上位論理デバイスの番号が登録される。エントリ７４には、当該外部デバイスに対応付けられ、当該外部デバイスに対するホスト１００からのアクセス要求を処理する、ストレージ１３０のＣＡの番号と、当該外部デバイスに対応付けられている下位論理デバイスの番号が格納される。また、ＡＡが直接当該外部デバイスを使用する場合は、ＣＡの番号ではなく、ＡＡの番号が格納される。なお、当該外部デバイスに対して上位論理デバイスもしくは下位論理デバイスが割り当てられていない、もしくはＡＡが直接使用しない場合、エントリ７３もしくはエントリ７４には無効値が設定される。

【００９０】

エントリ７５には、当該外部デバイスの状態を示す情報が設定される。各状態の種類は物理デバイス管理情報２０２のエントリ６４と同じである。ストレージ

１３０は初期状態では外部ストレージ１８０を接続していないため、エントリ７５の初期値は「未実装」となる。エントリ７６には、当該外部デバイスを有する外部ストレージ１８０の識別情報が登録される。外部ストレージ１８０の識別情報としては、外部ストレージ１８０のベンダ識別情報と各ベンダが一意に割り振る製造シリアル番号の組み合わせなどが考えられる。

【００９１】

エントリ７７には、当該外部デバイスに対して当該外部デバイスを有する外部ストレージ１８０が割り振ったデバイスの識別番号が格納される。なお、外部デバイスは外部ストレージ１８０の論理デバイスであるから、エントリ７７には外部ストレージ１８０における論理デバイス番号が格納される。

【００９２】

エントリ７８には、当該外部デバイスへアクセス可能なストレージ１３０のポート１４１に割り振られた識別番号と、当該ポート１４１が属するＰＡの識別番号が登録される。複数のポート１４１から当該外部デバイスへアクセスできる場合には、複数のポート識別番号と複数のＰＡ識別番号が登録される。

【００９３】

エントリ７９には、当該外部デバイスが外部ストレージ１８０の１つ以上のポート１８１にＬＵＮ定義されている場合、それらのポート１８１のポートＩＤおよび当該外部デバイスに割り当てられたターゲットＩＤ／ＬＵＮが１つ又は複数個保持される。なお、ストレージ１３０のＰＡが外部デバイスにアクセスする場合（ＰＡから外部デバイスに対する入出力要求を送信する場合）には、当該外部デバイスが属する外部ストレージ１８０によって当該外部デバイスに割り当てられたターゲットID及びLUNが、当該外部デバイスを識別するための情報として用いられる。

【００９４】

図１６は、NAS管理情報２２３の一構成例を示す図である。NAS構成情報２２３は、各ＡＡごとに、ＡＡ番号～代替ＡＡ番号の情報の組を登録するエントリ１６０１～１６０６を有する。

エントリ１６０１には、ストレージ１３０内でＡＡを識別するための識別情報

が登録される。

【００９５】

エントリ１６０２には、ホスト１００がファイルアクセス用の通信コマンドを送信する際に送信先を指定するためのIPアドレス、当該ＡＡが管理するファイルが存在するディレクトリ名、当該ディレクトリ名で指定されるディレクトリの情報とディレクトリ下に存在するファイルを格納するために割り当てられている上位論理デバイスの番号もしくは外部デバイス番号、及び当該上位論理デバイスをＡＡに対して提供しているＣＡの番号の組からなる情報のリストが登録される。なお、IPアドレスのみを登録し、登録されたIPアドレス先に来た全要求を当該ＡＡが処理するようにしても良い。

【００９６】

エントリ１６０２に複数のIPアドレスが登録されている場合、ホスト１００からは、複数のNASがシステム内にあるかのように見える。

また、１つのIPアドレスに複数のＡＡを割り当てて、例えばディレクトリ毎に異なるＡＡが処理するようにすることもエントリ１６０２の設定で可能である。また、ＡＡが複数の上位論理デバイスを例えばディレクトリ毎に使い分ける場合もエントリ１６０２の設定で可能である。

【００９７】

エントリ１６０３には、ＡＡの処理状態を示す情報が登録される。登録される情報には、「オフライン」、「障害」及び「オン中」の３つが有る。

「オフライン」とは、ＡＡが初期設定処理中だったり、他ＡＡからの情報引継ぎ処理中であってホスト１００からの要求を受け付けない状態を示す。

「障害」とは、ＡＡに障害が発生した状態を示す。

「オン中」とは、ＡＡがホスト１００からの要求を受け付け可能な状態を示す。

【００９８】

エントリ１６０４には、ＡＡのアクセスモードを示す情報が登録される。アクセスモードを示す情報とはＡＡが他の部位（ＣＡやＰＡ）と連携して動作を行うかどうかを示す情報であり、「連携」と「非連携」の二種類がある。

「非連携」が設定される時には、ＡＡの処理にかかわるコマンドおよびデータは、全てＡＡを経由する。具体的には、ホスト１００からＰＡが受信したコマンドはＡＡに送信され、ＡＡは、必要に応じ入出力要求をＣＡもしくはＰＡ経由で外部ストレージ１８０に送信する。又、その応答コマンドおよびデータはＣＡもしくはＰＡからＡＡに送信され、処理終了後、ＡＡは応答コマンドをホスト１００からのコマンドを受信したＰＡに送信する。

【００９９】

「連携」が設定される時には、ＡＡの入出力要求に基づく、ＣＡ又はＰＡ経由で接続された外部ストレージ１８０から送信されるデータはＡＡを介さず直接ホスト１００からコマンドを受信したＰＡへ送信され、ＣＡ又はＰＡ経由で接続された外部ストレージ１８０へ送信するデータは、ホスト１００からコマンドを受信したＰＡから直接ＣＡ又は外部ストレージ１８０と接続されているＰＡに対し送信される。これにより、大量のデータがＡＡを経由しないため、ＡＡの処理負荷と相互接続網１７０の負荷が軽減される。なお、本実施形態では、簡単のため、ストレージ１３０の工場出荷時にあらかじめ、初期値として「非連携」がエントリ１６０４に設定されており、その後、適宜、ストレージ管理者または、管理サーバからの制御により後述の手段を用いて値が設定変更されるものとする。「連携」は、例えばサイズの大きいファイルに対するアクセスが多いと推測される環境において、処理負荷の軽減のためにストレージ管理者により採用される。なお、エントリ１６０４を省略し、ＡＡとして「連携」もしくは「非連携」のいずれか一方を提供することにしてもよい。

【０１００】

エントリ１６０６には、当該ＡＡに障害が発生した場合に、当該ＡＡの管理情報を引き継ぎ処理を行う代替ＡＡの番号が設定される。代替ＡＡが無い場合には、無効値が設定される。代替ＡＡを設けない場合には、エントリ１６０６は省略されてもよい。

【０１０１】

図１７は、ポート管理情報２２４の一構成例を示す図である。ポート管理情報２２４は、ポート１４１がIPネットワークと接続される場合にその接続状態を管

理する際にＭＡ等の各部位で用いられる。尚、ポート１４１がFCネットワークに接続される場合には、そのポート１４１は、ポート管理情報２２４により管理されない。ポート管理情報２２４は、ＩＰネットワークと接続されるポート１４１ごとに、そのポートに関する情報を登録するエントリ１７０１～１７０８を有する。

【０１０２】

エントリ１７０１には、ポート１４１をストレージ１３０内で一意に識別するために割り付けれらた番号が設定される。エントリ１７０２には、当該ポート１４１のMACアドレスが登録される。エントリ１７０３には、当該ポート１４１に割り当てられたIPアドレスのリストが登録される。エントリ１７０４には、当該ポート１４１が処理できるプロトコル（例えば、iSCSI、NFS等）のリストが登録される。ポート１４１は、自ポートのエントリに登録されたMACアドレス１７０２、IPアドレス１７０３、プロトコル１７０４に従い通信処理を行う。なお、処理できないプロトコルのデータをポート１４１が受信した場合には、当該データは、ポート１４１にて廃棄処理される。

【０１０３】

エントリ１７０５には、当該ポート１４１の状態が登録される。状態には「通信可能」、「ネットワーク非接続」、「通信不可能」及び「障害」の四種類がある。

「通信可能」は、ポート１４１に障害が無くかつネットワークに接続されており、ホスト１００との通信が可能な状態にあることを示す。「ネットワーク非接続」は、ポート１４１にケーブル等が接続されておらずネットワークが接続されていない状態を示す。「通信不可能」は、ポート１４１がネットワークに接続されているが、ネットワークの問題でホスト１００と通信ができない状態にあることを示す。「障害」は、ポート１４１に障害ある状態を示す。

【０１０４】

エントリ１７０７には、ポート１４１が外部ストレージ１８０に接続する場合に、当該ポート１４１に接続された外部ストレージ１８０の外部デバイス番号が登録される。ポート１４１が外部ストレージ１８０に接続されていない場合には

、無効値が設定される。

【０１０５】

エントリ１７０８には、当該ポート１４１に障害が発生した場合に、当該ポート１４１に割り当てれらたIPアドレスのリスト１７０３を引継ぎ当該IPアドレスに対するアクセスを続行するポート１４１（以下「代替ポート」）のポート番号が登録される。なお、代替ポートが通信に使用されていなかった場合に引継ぎ処理を行う場合、IPアドレスのリストだけでなく、MACアドレスも代替ポートへ引き継いでもかまわない。

【０１０６】

なお、ポート管理情報２２４でFCプロトコルに対応するするポートも管理する場合、前述の情報に加え、エントリ１７０２にはWWNが、エントリ１７０３にはポートIDが、エントリ１７０４にはFCを示す情報が登録される。

【０１０７】

以下、ホスト１００がストレージ１３０に対してファイルアクセスを行った場合のストレージ１３０内部での処理手順について説明する。ファイルアクセスの処理は、具体的には、要求振り分け処理、ファイルアクセス処理、連携モード処理、コマンド処理及びライトアフタ処理の各処理から構成される。以下、順に説明する。

【０１０８】

図１１は、要求振り分け処理２５１の処理フローの一例を示す図である。要求振り分け処理２５１は、ＰＡのポート１４１においてホスト１００または外部ストレージ１８０から受け取った入出力要求やデータ、および転送制御部１４４においてＣＡ、ＡＡ又は別のＰＡから受け取った入出力要求やデータやコマンドを、ＣＡ、ＡＡ、他のＰＡ、ホスト１００又は外部ストレージ１８０等、適切な対象部位へＰＡが振り分ける処理である。

【０１０９】

ポート１４１又は転送制御部１４４からデータやコマンドなどを含んだフレームを受信したＰＡは（ステップ１１０１）、フレーム内のプロトコルの種別および当該プロトコルのコマンドの種別によって処理を切り分ける。なお、プロトコ

ルの種別は、当該ポート１４１がサポートするネットワーク（IPもしくはFC）と受信したフレーム内のデータから判別することができる（ステップ１１０２）。

【０１１０】

当該フレームがＦＣＰコマンドフレームの場合もしくはiSCSIコマンドであった場合、ＰＡは、当該フレームの送信元によって処理を切り分ける。なお、本実施形態では、コマンドフレームは、ポート１４１を介してホスト１００から受信する場合と、転送制御部１４４を介して外部ストレージ１８０をアクセスするＣＡ、ＡＡ又は他ＰＡから受信する場合の２つがある（ステップ１１０３）。

【０１１１】

まず、送信元がホスト１００の場合、ＰＡは、受信したコマンドのヘッダ情報に含まれるアクセス対象の上位論理デバイスを表すLUNに基づいて、ＬＵパス管理情報の複製２１６及び上位論理デバイス管理情報の複製２１４を参照し、当該上位論理デバイスが外部デバイスに対応付けられており、かつデバイスアクセスモードが「ＣＡ経由」でないかを判断する。当該上位論理デバイスが外部デバイスに対応付けられておりかつデバイスアクセスモードが「ＰＡ直結」である場合、ＰＡは更に、当該コマンドの要求種別がＣＡでの処理不要であるかを判断する。これらの条件を全て満たす場合、ＰＡは、当該要求のアクセス形態はＰＡ直結であると判断し、いずれかの条件を満たさない場合は当該要求のアクセス形態はＣＡ経由であると判断する（ステップ１１０４～１１０５）。

【０１１２】

ステップ１１０５でアクセスモードが「ＣＡ経由」でなく、コマンド要求種別がＣＡ処理不要の場合、ＰＡはさらにアクセスモードが「ＰＡ直結移行中」でかつ当該コマンド内ＬＢＡとブロック数で示されるアクセス領域が当該上位論理デバイスの切替進捗ポインタよりも後にあるかを判断する（ステップ１１１９）。アクセスモードが「ＰＡ直結移行中」で無い又はＰＡが受信したコマンドが進捗ポインタよりも前、すなわちＰＡ直結へ切替済みの領域を対象としている場合には、ＰＡは、コマンドのアクセス先を外部デバイスを有する外部ストレージ１８０が接続されたＰＡとする（ステップ１１０６）。

【０１１３】

ステップ１１１９でアクセスモードが「ＰＡ直結移行中」でかつアクセス先が未切替な部分であると判断された場合及びステップ１１０５でＣＡ経由アクセスであると判断した場合、ＰＡは上位論理デバイス管理情報の複製２１４を参照して、受信したコマンドのアクセス対象となっている上位論理デバイスが対応するＣＡ番号などを算出し、これを入出力処理の要求先、当該フレームの送信元を入出力処理の要求元として、当該コマンドのルーティング制御情報として記憶する（ステップ１１０７）。

【０１１４】

なお、ＰＡ直結移行中処理を行わない場合、ステップ１１１９は省略してもかまわない。この場合、図３のエントリ３９は省略される。

【０１１５】

一方、ステップ１１０３の切り分け処理で、当該フレームの送信元がホスト１００以外、すなわちＣＡ、ＡＡ又は他ＰＡ１４０の場合、ＰＡは当該フレームのヘッダ情報で指定されている外部ストレージ１８０の外部デバイスを転送先に決定し、これを当該入出力処理の要求先、当該フレームの送信元を入出力処理の要求元として当該コマンドのルーティング制御情報を登録する（ステップ１１０８、１１０９）。

【０１１６】

一方、ステップ１１０２の切り分け処理で、ＰＡ１４０で受信したフレームの種別がデータフレーム、受信準備完了通知フレーム又は完了報告フレームの場合、ＰＡは、当該フレームに対応する入出力処理のルーティング制御情報を参照し、当該フレームの送信元部位の交信相手先を当該フレームの転送先として決定する。すなわち、ＰＡが当該フレームを外部ストレージ１８０から受信した場合は、当該入出力処理の要求元であるホスト１００が接続されたＰＡ、ＣＡ又はＡＡがフレームの転送先となる。又、ＰＡが当該フレームをホスト１００から受信した場合は、当該入出力処理の要求先であるＣＡ又は外部ストレージ１８０が接続されたＰＡが当該フレームの転送先として決定される。さらに、ＰＡが当該フレームをＣＡ又は他ＰＡから受信した場合は、当該入出力処理の交信相手先（即ち当該フレームの送信元が当該入出力処理の要求元ならば要求先）であるホスト１

００又は外部ストレージ１８０が、それぞれ当該フレームの転送先として決定される（ステップ１１１０）。

【０１１７】

ステップ１１０６、１１０７、１１０９及び１１１０で受信フレームの転送先が決まったら、ＰＡは転送先部位によって処理を切り分ける（ステップ１１１１）。

転送先が他のＰＡの場合、ＰＡは、当該入出力要求のルーティング制御情報に登録されている転送先のアドレスを当該フレームの送信先アドレスへ設定する。すなわち、送信先が外部デバイスの場合、ＰＡは、受信したフレームのデスティネーションＩＤをルーティング制御情報に当該入出力処理の要求先として登録されている外部デバイスのターゲットポートID（即ちポート１８１のポートＩＤ）へ、また、同フレーム内のＬＵＮなどのアドレス情報を当該外部デバイスのアドレス（即ち外部デバイス管理情報を参照して取得したLUN等）に書き換える。送信先がホスト１００の場合、ＰＡは、受信フレームのデスティネーションＩＤをルーティング制御情報に当該入出力処理の要求元として登録されているホスト１００のポート１０７のポートＩＤへ書き換える（ステップ１１１２）。

【０１１８】

また、転送先が当該ＰＡ内のポート１４１に接続されたホスト１００又は外部ストレージ１８０の場合は、ＰＡは、当該ポート１４１のアドレスを当該フレームの送信元アドレスへ設定する（ステップ１１１３）。

そして、コマンド、データ、受信準備完了報告、完了報告フレームを受信したＰＡは、決定された転送先であるホスト１００、又は外部ストレージ１８０、ＣＡ、ＡＡ又は他のＰＡへ当該フレームを転送する（ステップ１１１４）。さらに、転送を完了した当該フレームが完了報告フレームだった場合は、ＰＡは、当該入出力処理のルーティング制御情報の登録を解除する(ステップ１１１５，１１１６)。

【０１１９】

また、ステップ１１０２で、ＰＡが受信したフレームがNASのコマンド、即ちファイルアクセスのコマンドであった場合、ＰＡは、フレームヘッダにセットさ

れているIPアドレスおよびコマンドにセットされているファイル名から得られるディレクトリ情報を元に、NAS管理情報の複製221を検索し、コマンドを処理するべきＡＡを決定し、当該ＡＡにコマンドを転送する。つぎに、ＡＡはファイルアクセス処理２２９（ステップ１１１８）を行う。尚、ファイルアクセス処理２２９の詳細については、後述する。

【０１２０】

また、ステップ１１０２で、ＰＡが受信したフレームがＦＣＰもしくはiSCSI関連のフレームでなく、かつ、NASのコマンドでもなかった場合は、フレームがＦＣのフレームであればファイバチャネルのノードポートとして従来公知な処理を当該フレーム受信ＰＡの制御プロセッサ１４２にて実行し、フレームがＩＰのフレームであればＰＡはフレームの廃棄処理を行う（ステップ１１１７）。

【０１２１】

なお、要求振り分け処理が実行された結果、コマンドの転送先がＣＡ又はＡＡであると決定された場合、ＰＡは、フレーム以外にアクセス対象であるＣＡ番号またはＡＡ番号、下位論理デバイス番号または上位論理デバイス番号または外部デバイス番号、ＬＢＡ、サイズなどの情報を付加して転送先のＣＡもしくはＡＡへ送信する。また、コマンドに続くデータなどをＣＡもしくはＡＡに転送する場合、ＰＡ、ＣＡ又はＡＡのそれぞれの転送制御部において、既に確立したコマンドシーケンスを意識したデータ転送を行う。

【０１２２】

図２１は、ファイルアクセス処理２２９の処理フローの一例を示す図である。要求振り分け処理２５１の結果、ＡＡはＰＡからコマンドを受信する（ステップ２１０１）。ＡＡは、コマンドで要求された処理対象のファイル関連の情報が自身のメモリ１４９にキャッシュされているかをキャッシュ管理情報２３０でチェックする（ステップ２１０２）。処理対象のファイル関連の情報（ディレクトリ情報等）が無かった場合、ＡＡはNAS管理情報の複製２２７を参照して関連情報が記憶されているデバイス（ＣＡが提供する論理デバイス又は外部デバイス）からディレクトリ情報を読み出しメモリ１４９へ格納する（ステップ２１０３）。

【０１２３】

情報があった場合、または、情報を読み出しキャッシュに格納した後、ＡＡは、要求がリードもしくはライトであるか否かをチェックする（ステップ２１０４）。リード又はライトでなかった場合、ＡＡは既存技術を使い当該コマンドの処理を行う（ステップ２１１８）。

【０１２４】

要求がリード又はライトであった場合、ＡＡはデータが自己のメモリ１４９のキャッシュ領域にあるか否かチェックする（ステップ２１０５）。要求がリードで、かつ既にデータがメモリ１４９のキャッシュ領域にあった場合、ＡＡはそのデータを元にリード処理を行う。

要求がリードで、かつデータがメモリ１４９のキャッシュ領域に無かった場合、ＡＡはメモリ１４９にキャッシュ領域を確保する（ステップ２１０６）。キャッシュ領域を確保後、もしくは、要求がライトであった場合、ＡＡ自身のアクセスモードが「連携」か否かをチェックする（ステップ２１０７）。

【０１２５】

「連携」でなかった場合、ＡＡはコマンドの内容とディレクトリ情報に従い、当該データが記録されているＣＡ又は外部ストレージ１８０が接続されているポート１４１があるＰＡへ入出力要求を送信する（ステップ２１０８）。送信したコマンドがリード要求であった場合（ステップ２１０９）、ＡＡは、コマンド処理２５４を実施したＣＡもしくは要求ふり分け処理２５１を実施したＰＡからデータを受信し、受信したデータをキャッシュ領域に格納し、その後完了報告を受信（ステップ２１１０、２１１３）する。

【０１２６】

ステップ２１０９で送信したコマンドがライト要求であった場合、ＡＡは、コマンド処理２５４を実施したＣＡもしくは要求ふり分け処理２５１を実施したＰＡから準備完了報告を受信し（ステップ２１１１）、ついでデータを送信し（ステップ２１１２）、完了報告を受信する（ステップ２１１３）。

その後、ＡＡは受信したデータに基づいて、コマンドで要求された操作ができる分のデータのリードを完了したか、またはデータのライトを完了したかをチェックし、完了していなかった場合、ＡＡはステップ２１０１～２１１４の処理を

データのリードまたはライトが完了するまで繰り返す（ステップ２１１４）。

ステップ２１０７で連携モードであった場合、ＡＡ、ＰＡ、ＳＡは、連携モード処理を行う（ステップ２１１７）。

【０１２７】

ステップ２１０５で処理を完了した場合、ステップ２１１４、ステップ２１１７及びステップ２１１８の処理の終了後、ＡＡは、その結果をコマンド応答にしてコマンド送信元のＰＡへ送信する(ステップ２１１５)。コマンド応答を受信したＰＡは、コマンド送信元のホスト１００へ当該コマンド応答を送信する（ステップ２１１６）。

【０１２８】

なお、本実施形態では、ＰＡでは、TCP/IPまでの処理を行い、NFS等NASのアクセスに使用されるプロトコルにしたがったコマンドがＰＡとＡＡ間で送受される。

また、本実施形態では、ライト要求に伴うデータを直接ＣＡに書き込む形をとったが、一旦ＡＡのメモリ１４９のキャッシュ領域に書き込んでから、適当な時にＣＡに書き込む方式でも良い。また、本実施形態では、ＡＡが、1回以上のリード／ライト要求をＣＡ又は外部ストレージ１８０に接続されるＰＡに対して行う方式をとったが、ＣＡ、ＰＡ及びＡＡとの間で指示方法を予め決めておき、一回のリード／ライト要求で処理が終了するような方式としても良い。

【０１２９】

図２２は、ファイルアクセス処理２２９の処理内で実施される連携モード処理２１１７の処理フローの一例を示す図である。連携処理の場合、ＡＡは、コマンド送信元のＰＡへ連携モードで動作することを通知する（ステップ２２０１）。次にＡＡは、リードまたはライトを行うデバイスの処理を行うＣＡまたは外部ストレージ１８０と接続されるポート１４１があるＰＡへ連携モードでのリード／ライト要求を送信する（ステップ２２０２）。

【０１３０】

送信したのがリード要求であった場合（ステップ２２０３）、コマンド送信元のＰＡは、コマンド処理２５４を実施したＣＡ又は要求ふり分け処理２５１を実

施したＰＡからデータを受信（ステップ２２０４）し、ＡＡは完了報告をコマンド処理２５４を実施したＣＡ又は要求ふり分け処理２５１を実施したＰＡから受信（ステップ２２０５）する。ステップ２２０３でライト要求であった場合、コマンド送信元のＰＡは、コマンド処理２５４を実施したＣＡ又は要求ふり分け処理２５１を実施したＰＡから準備完了を受信し（ステップ２２０７）、ついでデータを送信する（ステップ２２０８）。その後、ＡＡは、その完了報告をコマンド処理２５４を実施したＣＡ又は要求ふり分け処理２５１を実施したＰＡから受信する（ステップ２２０５）。

【０１３１】

ＡＡは受信したデータでコマンドで要求された操作ができる分のデータのリードを完了したか、またはデータのライトを完了したかをチェックし（ステップ２２０６）、完了していなかった場合は、ステップ２２０２～２２０６の処理をデータのリードまたはライトが完了するまで繰り返す。

【０１３２】

データリードまたはライトが完了した後、ＡＡは図２１の処理フローに戻り、結果をコマンド応答にしてコマンド送信元のＰＡへ送信し(ステップ２１１５)、ＰＡは、先にＡＡから受信した連携モード通知、ＰＡまたはＣＡから受信したデータ、ＡＡから受信したコマンド応答からコマンド応答を作成し、コマンド送信元のホスト１００へ当該コマンド応答を送信する（ステップ２１１６）。

【０１３３】

図１２は、コマンド処理２５４の処理フローの一例を示す図である。コマンド処理は、リードやライトの要求を受信したＣＡにおいて下位論理デバイスに対する入出力要求を処理するものである。

【０１３４】

要求振り分け処理２５１の結果又はファイルアクセス処理２２９の結果、ＰＡ又はＡＡから受信したコマンドについて、ＣＡは、コマンド種別をチェックする（ステップ１２０２）。当該コマンドの要求がリードの場合、ＣＡは、ディスクキャッシュ１５４上にリード対象のデータがヒットしているか（キャッシュに格納されているか）をキャッシュ管理情報２０３を参照して判定する（ステップ１

２０３）。

【０１３５】

キャッシュミスの場合、ＣＡはキャッシュ管理情報２０３を更新してディスクキャッシュ１５４に領域を確保し（ステップ１２０４）、当該データが格納される物理デバイス又は外部デバイスからキャッシュへデータを読み出す。リードデータが格納されているデバイスが物理デバイスであるか外部デバイスであるかは、ＣＡが下位論理デバイス管理情報２１０を参照することによって判断することができる（ステップ１２０５）。

【０１３６】

リードデータが格納されているデバイスが物理デバイスであれば、ＣＡは物理デバイス管理情報２０２より特定したディスク装置１５７に対してリード要求を発行してデータを読み出し（ステップ１２０６）、読み出したデータをディスクキャッシュ１５４の確保領域に格納する（ステップ１２０７）。

【０１３７】

リードデータが格納されているデバイスが外部デバイスであれば、ＣＡは外部デバイス管理情報２０５を参照してエントリ７８からアクセス対象ＰＡを特定し、外部デバイスに対するリード要求を当該ＰＡへ送信する（ステップ１２１０）。

リード要求を受信したＰＡでは要求振り分け処理２５１によって当該リード要求を外部ストレージ１８０に対して送信し、応答として外部ストレージ１８０からリードデータを受信し、外部ストレージ１８０から受信したリードデータをＣＡへ返送する（ステップ１２１１）。ＣＡは、ＰＡから受信したリードデータをディスクキャッシュ１５４の確保領域に格納する（ステップ１２０７）。

【０１３８】

ディスクキャッシュにデータが存在していた又はディスクキャッシュにリードデータが格納されたら、ＣＡは、ディスクキャッシュ１５４に格納されたデータをコマンド転送元のＰＡ又はＡＡへ送信し（ステップ１２０８）、ＰＡは要求振り分け処理２５１を実行して当該データをホスト１１０へ転送し、ＡＡはファイルアクセス処理の続きを行う。

【０１３９】

ＣＡがＰＡ又はＡＡから受信したコマンドがライトの場合、ＣＡはライトデータに対応する旧データがディスクキャッシュ１５４に格納されているか否かを判断し（１２１２）、キャッシュミスの場合にはディスクキャッシュ１５４領域を割り当てる（１２１３）。

【０１４０】

ステップ１２１２又は１２１３の終了後、ＣＡは、コマンド転送元のＰＡ又はＡＡに対して、転送準備完了を送信する（ステップ１２１４）。ＰＡへ転送準備完了を送信した場合、転送準備完了を受信したＰＡは要求振り分け処理２５１を実行して転送準備完了をホスト１００へ転送し、その結果ホスト１００からライトデータが送られてくると、ＰＡは転送準備完了を送信したＣＡへデータを送信する。一方、ＡＡへ転送準備完了を送信した場合、転送準備完了を受信したＡＡは、ＣＡへデータを送信する。ＣＡではＰＡ又はＡＡから受信したライトデータを、確保したディスクキャッシュ１５４上の領域へ格納する（ステップ１２１５、１２１６）。

【０１４１】

また、ＣＡで受信したコマンドがリードでもライトでもない場合、例えばモードセンスなどのセンス系コマンドや診断系のコマンドの場合は、ＣＡは、一般的なＳＣＳＩストレージの仕様に基づき適切な処理を実行する（ステップ１２１７）。

【０１４２】

ステップ１２０８、１２１７又は１２１６の処理を終了したＣＡは、コマンド処理を完了したことをコマンド転送元のホスト１００に接続されているＰＡ又はＡＡへ報告する。完了報告を受領したＰＡは、要求振り分け処理２５１を実行して、ホスト１００へ完了報告フレームを送信する。完了報告を受領したＡＡは、ファイルアクセス処理の続きを行う（ステップ１２０９）。

【０１４３】

なお、ステップ１２０１でＣＡがＡＡからコマンドを受信した場合でＡＡが「連携」モードで動作している場合、ＣＡは、ＡＡからはコマンドの他に連携モー

ドの通知と連携するべきＰＡを指定する情報等を受信する。この場合、ＣＡは、ステップ１２０８およびステップ１２１４、１２１５における準備完了とデータの送受信を、連携を指示されたＰＡに対して行う。「連携」モードで動作しているＰＡは、ＡＡからの指示とＣＡからの準備完了およびデータの送受信に従い、連携モード時のファイルアクセス処理を完了する。

【０１４４】

図１３は、ライトアフタ処理２５７の処理フローの一例を示す図である。ライトアフタ処理とは、ＣＡにおけるコマンド処理２５４の結果、ディスクキャッシュ１５４に格納されたライトデータをＣＡがディスク装置１５７へ書き出す処理である。ディスクキャッシュ１５４上に保持されたライトデータはキャッシュ管理情報２０３で管理される。通常、ライトデータやディスクからリードされたリードデータは、なるべく古いものから順にディスクキャッシュ１５４より追い出されるように、キューなどで管理されている。

【０１４５】

ＣＡは、キャッシュ管理情報２０３で管理されたデータの中から、ディスク装置１５７へ実際に書き込むデータを選択し（ステップ１３０１）、当該データが格納されるディスク装置１５７が外部デバイスに対応づけられているか物理デバイスに対応付けられているかを下位論理デバイス管理情報２０１に基づいて判断する（ステップ１３０２）。

【０１４６】

ライトデータを物理デバイスに書き込む場合、ＣＡは物理デバイス管理情報２０２を参照してデータを書き込むディスク装置１５７を特定し、ライトデータを特定されたディスク装置１５７へ転送する（ステップ１３０３）。

【０１４７】

一方ライトデータを外部デバイスに書き込む場合には、ＣＡは外部デバイス管理情報２０５を参照してライトデータが書き込まれる外部デバイスを有する外部ストレージ１８０と接続されているＰＡを特定し、当該ＰＡへライト要求を送信する（ステップ１３０４）。ライト要求を受信したＰＡは、要求振り分け処理２５１を実行して当該ライト要求を外部ストレージ１８０に転送し、ライト要求に

対する応答として当該外部ストレージ１８０から転送準備完了通知を受信する。すると当該ＰＡは転送準備完了通知をライト要求送信元のＣＡへ送信する。完了通知を受信したＣＡはライトデータを当該ＰＡへ送信する（ステップ１３０５、１３０６）。

【０１４８】

ライトデータを受信したＰＡは、要求振り分け処理２５１を実行してライトデータを上述の外部ストレージ１８０に送信し、外部ストレージ１８０によってライトデータが外部デバイスに書き込まれる。

【０１４９】

ライトデータを物理デバイスもしくは外部デバイスへライトした後、ＣＡは、ライトデータが格納されていたディスクキャッシュ１５４上の領域を解放する（ステップ１３０７）。

【０１５０】

次に、ＡＡ又はＰＡのポート１４１に障害が発生した場合のストレージ１３０の動作について説明する。ＡＡ又はＰＡのポート１４１に障害が発生した場合、ストレージ１３０は、代替のＡＡ又はポート１４１を用いて処理を継続する。

【０１５１】

図２０は、NAS／ポート障害処理２３１の処理フローの一例を示す図である。ポート１４１もしくはＡＡの障害を検出または他の部位から通知されたＭＡは、検出もしくは通知された障害がポート障害かＡＡ障害かを判断する（ステップ２００１、２００２）。検出等された障害がポート障害であった場合、ＭＡは、障害が発生したポート１４１のポート管理情報２２４を参照し、当該ポート１４１の管理情報のエントリ１７０５を「障害」に設定する。

【０１５２】

次に、ＭＡは、エントリ１７０８に登録されている代替ポートのポート管理情報に障害の発生したポート１４１の情報を追加コピーする。代替ポートのエントリ１７０５が「通信」状態の場合は、障害の発生したポート１４１のエントリ１７０４に登録されていたIPアドレスリストを代替ポートのIPアドレスリストに追加する。「通信」状態でなかった場合、ＭＡは、障害の発生したポート１４１の

MACアドレス、IPアドレスリスト、アクセスモード等を代替ポートに対応するポート管理情報の各エントリにコピーする。そして、ＭＡはＰＡにポート管理情報の更新を指示する（ステップ２００３）。

【０１５３】

ＡＡ障害であった場合、ＭＡは障害が発生したＡＡに対応するNAS管理情報２２３を参照し、当該管理情報２２３のエントリ１６０３を「障害」に設定する。次に、ＭＡは障害が発生したＡＡに関するNAS管理情報２２３のエントリ１６０６に登録されている代替ＡＡのNAS管理情報２２３に、障害の発生したＡＡのIPアドレス、ディレクトリ名、対応CA番号、上位論理デバイス／外部デバイスのリストを追加する。その後、ＭＡは、ＡＡとＰＡにNAS管理情報２２３の更新を指示する（２００４）。

【０１５４】

ＭＡ、ＰＡ及びＡＡでポート管理情報とNAS管理情報の更新が終了した後、ＭＡは、ＳＴにポート管理情報とNAS管理情報の更新を通知する（ステップ２００５）。ポート管理情報とNAS管理情報の更新を通知されたＳＴは、更新のあった情報を取り込んだ後、要求元であるストレージ管理者もしくは管理サーバへ報告すべく完了報告を出力する（ステップ２００６）。

【０１５５】

図１８は、NAS管理情報変更処理２２５の処理フローの一例を示す図である。なお、新規にNAS管理情報を登録する場合、および、あるＡＡにかかわるNAS管理情報の一部もしくは全てを別ＡＡにコピーして処理を引き継がせる場合も本処理が実行される。更に本処理は、ＡＡおよびＰＡが各自の処理負荷を測定する機構を持ち、その処理負荷の情報をＭＡもしくはＳＴ経由で管理サーバが収集し、その情報を元に、ストレージ管理者または、管理サーバ上の記載しない処理負荷管理処理部が処理負荷の高いＡＡもしくはＡＡの処理にかかわるＰＡから、処理負荷の低いＡＡもしくはＰＡへ負荷を分散させるように制御する時に動作する。

【０１５６】

また、一部のＡＡの修理や、ＡＡを高性能なものに変更する等のために新しい物と交換するような場合、ストレージ管理者または、管理サーバからの制御で動

作させる。

【０１５７】

ストレージ管理者もしくは管理サーバ１１０からNAS管理情報変更指示を受け付けたＳＴは、ＭＡに対して同指示を転送する（ステップ１８０１）。同指示には、情報を変更するＡＡにかかわるNAS管理情報エントリが付加される。指示を受信したＭＡは、NAS管理情報を変更すると同時に、変更するNAS管理情報のエントリにかかわるＡＡに対し、NAS管理情報を変更するように指示を送信する（ステップ１８０２）。指示を受信したＡＡは、NAS管理情報の複製２２７の変更を行う（ステップ１８０３）。変更指示が新たなエントリの追加であった場合（ステップ１８０４）、ＭＡに登録の完了を通知する（ステップ１８０７）。

【０１５８】

変更指示が既存エントリの変更であった場合、ＡＡは、当該変更にかかわる処理中のコマンドを廃棄し（ステップ１８０５）さらに、メモリ１４９にキャッシュされている当該変更にかかわるディレクトリ情報をクリアして（ステップ１８０６）、ＭＡに登録の完了を通知する（ステップ１８０７）。ＡＡからの登録完了通知を受信したＭＡは、NAS管理情報の更新があったことを各ＰＡ及びＳＴに通知する（ステップ１８０８）。

【０１５９】

NAS管理情報更新を通知されたＰＡでは、更新のあったNAS管理情報の複製２２１をメモリ１４３へ取り込む。又、NAS管理情報更新を通知されたＳＴでは、更新のあった情報を取り込んだ後、要求元であるストレージ管理者もしくは管理サーバ１１０へ報告すべく完了報告を出力する（ステップ１８０９）。

【０１６０】

尚、ＮＡＳ／ポート障害処理にてNAS管理情報が変更される際には、上述した処理手順のうち、ステップ１８０２から処理が開始される。

【０１６１】

図１９は、ポート管理情報変更処理２２６の処理フローの一例を示す図である。なお、新規にポート管理情報を登録する場合及びあるポート１４１にかかわるポート管理情報の一部もしくは全てを別ポート１４１にコピーして処理を引き継

がせる場合も、本処理が実行される。更に本処理は、ＰＡが各自の処理負荷を測定する機構を持ち、その処理負荷の情報をＭＡもしくはＳＴ経由で管理サーバが収集し、その情報を元に、ストレージ管理者または、管理サーバ上の記載しない処理負荷管理処理部が処理負荷の高いポートから、処理負荷の低いポートへ負荷を分散させるように制御する時に動作する。また、一部のＰＡの修理や、ＰＡを高性能なものに変更する等のために新しい物と交換するような場合、ストレージ管理者または、管理サーバからの制御で動作させる。

【０１６２】

ストレージ管理者もしくは管理サーバ１１０からNAS管理情報変更指示を受け付けたＳＴがＭＡに対して同指示を転送する。同指示には、情報を変更するポートにかかわるポート管理情報のエントリが付加される（ステップ１９０１）。指示を受信したＭＡは、ポート管理情報を変更すると同時に、変更するポート管理情報のエントリにかかわるポート１４１が存在するＰＡに対し、ポート管理情報の複製２２２を変更するように指示を送信する（ステップ１９０２）。

【０１６３】

指示を受信したＰＡは、ポート管理情報の複製２２２の変更を行う（ステップ１９０３）。変更指示が新たなエントリの追加または、既存エントリへの情報の付加であった場合（ステップ１９０４）、ＰＡはＭＡに登録の完了を通知し、更新後のポート管理情報の複製２２２の内容に従い自ＰＡ内のポート１４１で通信処理を開始する（ステップ１９０７）。変更指示が既存エントリの内容の変更であった場合、ＰＡは当該変更にかかわる処理をクリアし、更新後のポート管理情報の複製２２２の内容に従い自ＰＡ内のポート１４１で通信処理を開始した（ステップ１９０５）後、ＭＡに登録の完了を通知する（ステップ１９０７）。ＰＡからの登録完了通知を受信したＭＡは、ポート管理情報の更新があったことをＳＴに通知する（ステップ１９０８）。

【０１６４】

ポート管理情報更新を通知されたＳＴは、更新のあった情報を取り込んだ後、要求元であるストレージ管理者もしくは管理サーバ１１０へ報告すべく完了報告を出力する（ステップ１９０９）。

【０１６５】

尚、ＮＡＳ／ポート障害処理にてポート管理情報が変更される際には、上述した処理手順のうち、ステップ１９０２から処理が開始される。

【０１６６】

次に、本システムの各部位で動作するストレージ制御処理について説明する。ストレージ１３０内に存在する物理デバイス、及び外部ストレージ１８０に存在する外部デバイスを含むデバイスを、特定のホスト１００またはＡＡに割り当てて当該ホスト１００またはＡＡから使用できるようにする処理は、外部デバイス定義処理２５３、論理デバイス定義処理２５５、ＬＵパス定義処理２５２の大きく３つに分けることができる。システム管理者は、上述の３つの処理を順に行うことで、上位論理デバイスをホスト１００又はＡＡに割り当てるようにストレージ１３０を設定することができる。以下、この３つの処理について説明する。

【０１６７】

図８は、外部デバイス定義処理２５３の処理フローの一例を示す図である。外部デバイス定義処理２５３は、ストレージ１３０の管理下に外部ストレージ１８０内のデバイスを外部デバイスとして導入するための処理である。

【０１６８】

まず、ストレージ管理者もしくは管理サーバ１１０からの外部ストレージ１８０接続指示を受け付けたＳＴが、ＭＡに接続指示を送信する。接続指示には、接続対象となる外部ストレージ１８０を特定する情報、例えば、外部ストレージ１８０のポート１８１のWWN、IPアドレス、又は外部ストレージデバイスへＩｎｑｕｉｒｙコマンドを送信した際にその応答により得られる装置識別情報もしくはその両方の情報及び外部ストレージ１８０に接続されるストレージ１３０のポート１４１の番号とが付加される（ステップ８０１）。

【０１６９】

外部ストレージの接続指示を受信したＭＡは、接続指示に付加された全てのポート１４１の番号に対応する全ＰＡへ、外部ストレージ１８０を識別するための情報と外部ストレージ１８０と接続されるポート１４１の番号とを有する外部ストレージ接続指示を送信する（ステップ８０２）。

【０１７０】

ＭＡから外部ストレージ接続指示を受信したＰＡは、接続指示に付加された外部ストレージ１８０の識別情報（以下「外部ストレージ識別情報」）を用いて、接続対象となる外部デバイスを探索する。具体的には、外部ストレージ識別情報としてポート１８１のWWNまたはIPアドレスを受信した場合、ＰＡは接続指示に付加されたポート番号により指定されたポート１４１から外部ストレージ１８０のポート１８１に対応する全ＬＵＮに対してＩｎｑｕｉｒｙコマンドを送信し、Ｉｎｑｕｉｒｙコマンドに対する正常な応答を返したＬＵＮを外部デバイス登録候補とする。

【０１７１】

外部ストレージ識別情報として装置識別情報しか受信しない場合、ＰＡは、当該ＰＡの全ポート１４１各々によって検出された外部ストレージ１８０のノードポート（ノードポートログイン時に検出済み）のそれぞれに対応する全ＬＵＮにＩｎｑｕｉｒｙコマンドを送信する。その後、Ｉｎｑｕｉｒｙコマンドに対する正常な応答を返したデバイス（ＬＵＮ）について、ＰＡは応答内に含まれる装置識別情報を外部ストレージ識別情報と比較し、一致するデバイスを外部デバイス登録候補とする（ステップ８０３）。

【０１７２】

ＰＡは、検出された外部デバイス登録候補についての情報リストをＭＡに返却する。この情報リストには、各外部デバイス登録候補について外部デバイス管理情報２０５を設定するのに必要な情報が含まれる（ステップ８０４）。

【０１７３】

情報リストを受取ったＭＡは、受け取った外部デバイスリストに含まれる外部デバイスに採番し、外部デバイス管理情報２０５へデバイス情報を登録し、同情報に更新のあった旨を他の各部位へ通知する。なお、本処理ステップにおいてＭＡが外部デバイス管理情報２０５へ登録する情報には、具体的には、ＭＡが外部デバイスに対して割り当てた外部デバイス番号、Ｉｎｑｕｉｒｙコマンドに対する応答から取得されるサイズ、ストレージ識別情報及び外部ストレージ内デバイス番号、ＰＡからＭＡに通知されるＰＡ/イニシエータポート番号リスト及びタ

ーゲットポートＩＤ/ターゲットＩＤ/ＬＵＮリストが含まれる。

【０１７４】

また、対応上位論理デバイス番号及び対応ＣＡ番号/下位論理デバイス番号/ＡＡ番号は未割当のため、ＭＡは初期値である無効値を対応するエントリに設定する。更にエントリ７５には「オフライン」状態が設定される（ステップ８０５）。

【０１７５】

外部デバイス管理情報２０５の更新通知を受けた各ＣＡ、ＰＡ及びＡＡは、ＭＡの制御メモリ１６４に保存されている外部デバイス管理情報２０５を各々のメモリに読み込む。また、ＳＴは外部デバイス管理情報２０５をメモリに取り込むと共に、外部デバイス定義処理の完了を、当該処理の要求元であるストレージ管理者もしくは管理サーバ１１０に通知すべく完了通知を出力する（ステップ８０６）。

【０１７６】

なお、ストレージ１３０に対してストレージ管理者もしくは管理サーバ１１０が接続指示と共に、導入対象となる外部ストレージ１８０を指示せず、外部ストレージ１８０の接続指示のみをストレージ１３０に指示し、ストレージ１３０で全ポート１４１から検出した全外部ストレージ１８０各々が有する全デバイスを外部デバイスとして登録しても構わない。また、ストレージ管理者や管理サーバ１１０からは特に明示的な接続指示を与えず、ストレージ１３０に外部ストレージ１８０が接続されたのを契機に、ストレージ１３０が検出した全デバイスを外部デバイスとして登録してもよい。

【０１７７】

図９は、論理デバイス定義処理２５５の処理フローの一例を示す図である。論理デバイス定義処理２５５は、ストレージ管理者もしくは管理サーバ１１０からの指示を受けて、ストレージ１３０で定義される物理デバイス又は外部デバイス定義処理２５３で定義された外部デバイスに対して、上位又は下位論理デバイスを定義する処理である。

【０１７８】

ストレージ管理者もしくは管理サーバ１１０から論理デバイス定義指示を受け付けたＳＴは、当該指示をＭＡに送信する。論理デバイス定義指示には、論理デバイスが割り当てられる対象となる物理デバイス又は外部デバイスの番号、割り当てられる上位論理デバイスの番号、割り当てられる下位論理デバイスの番号及び当該下位論理デバイスを管理するＣＡの番号とが付加される。なお本実施形態では、説明の簡略化のため、１つの物理もしくは外部デバイスに対して１つの論理デバイスを割り当てるものとするが、２つ以上の物理もしくは外部デバイスからなるデバイスグループに対して１つの論理デバイスを、又は１つの物理もしくは外部デバイスに対して２つ以上の論理デバイスを、又は２つ以上の物理もしくは外部デバイスからなるデバイスグループに対して２つ以上の論理デバイスを定義しても構わない。

【０１７９】

ただし、それぞれの場合、下位論理デバイス管理情報２０１に、物理もしくは外部デバイス内での当該論理デバイスの開始位置およびサイズなどの情報の付加、対応する物理および外部デバイス番号エントリの複数デバイス化、物理および外部デバイス管理情報２０２、２０５に、対応する下位論理デバイス番号エントリの複数デバイス化を行うことがそれぞれ必要となる（ステップ９０１）。

【０１８０】

論理デバイス定義指示を受信したＭＡは、論理デバイス定義指示に付加されている情報より対象ＣＡを特定し、特定されたＣＡに論理デバイス定義指示を送信する（ステップ９０２）。

【０１８１】

論理デバイス定義指示を受信したＣＡは、指示された物理もしくは外部デバイスに対して下位論理デバイスを登録する（ステップ９０３）。具体的にＣＡは、下位論理デバイス管理情報２０１の対象デバイスエントリ（論理デバイス定義指示に付加されている下位論理デバイスの番号に対応するエントリ）について、論理デバイス定義指示に付加されている下位論理デバイス番号をエントリ５１に、論理デバイス定義指示に付加されている物理もしくは外部デバイスの番号をエントリ５３に、下位論理デバイスのサイズをエントリ５２に、論理デバイス定義指

示に付加されている上位論理デバイス番号をエントリ５５に各々登録し、エントリ５４には「オンライン」を設定する。

【０１８２】

また、当該物理／外部デバイスの対応ＣＡ番号／下位論理デバイス番号を設定し、デバイス状態を「オンライン」に更新する。登録が完了したらＣＡはＭＡにその旨通知する（ステップ９０３）。

【０１８３】

次にＭＡは、論理デバイス定義指示によって指定された下位論理デバイスに上位論理デバイスを割り当て、制御情報が更新された旨を他の各部位に通知する。具体的にＭＡは、上位論理デバイス管理情報２０３の当該デバイスエントリ（論理デバイス定義指示に付加されている上位論理デバイスの番号に対応するエントリ）について、論理デバイス定義指示に付加されている上位論理デバイスの番号をエントリ３１に、上位論理デバイスのサイズをエントリ３２に、論理デバイス定義指示に付加されているＣＡ番号と下位論理デバイス番号をエントリ３３に設定し、デバイス状態としてエントリ３４には「オフライン」状態を設定し、エントリ３５及び３６に関しては未割当のため無効値を設定する。

【０１８４】

更に、当該上位論理デバイスの割当対象が外部デバイスの場合には、論理デバイス定義指示に付加されている外部デバイス番号をエントリ３８に設定し、デバイスアクセスモードを初期値として「ＰＡ直結」に設定する。更に、当該上位論理デバイスの割り当て対象が外部デバイスである場合には、外部デバイス管理情報２０５の対応エントリ７３に当該上位論理デバイス番号を設定する。そして、ＭＡは制御情報の更新があった旨を各ＰＡ及びＳＴに通知する（ステップ９０４）。

【０１８５】

制御情報更新を通知されたＰＡは更新のあった情報をメモリ１４３へ取り込む。又、制御情報更新を通知されたＳＴは、更新のあった情報を取り込んだ後、論理デバイス定義処理の完了を要求元であるストレージ管理者若しくは管理サーバ１１０へ報告すべく終了報告を出力する（ステップ９０５）。

【０１８６】

図１０は、ＬＵパス定義処理２５２の処理フローの一例を示す図である。ストレージ管理者もしくは管理サーバ１１０からＬＵパス定義指示を受け付けたＳＴがＭＡに対して同指示を転送する。同指示には、対象上位論理デバイス番号、ＬＵを定義するポート１４１の番号、ＬＵＮに加えて、同ＬＵにアクセスするホスト１００の識別情報（ホストが有するポート１０７のＷＷＮなど）またはＡＡの識別情報（ＡＡ番号等）及び当該上位論理デバイスへの要求機能レベルを示す情報が付加される。

【０１８７】

ここで要求機能レベルとは、上位論理デバイスに要求されるデータの信頼性等のレベルを指す。具体的には、ストレージ１３０が提供するデータ複製やデバイス再配置などのデータ連携機能や、データをキャッシュ上に格納しておくキャッシュ常駐化によるアクセス高速化機能などの適用有無が指定される。これによって、機能の適用の有無により、データの信頼性のレベルが変化する。なお、要求機能レベルとして、単に、当該上位論理デバイスに格納されるデータがＣＡ経由で処理が必要か否か（たとえ、ＣＡに接続される記憶装置１５７にデータが格納されなくても）を示す情報を指定してもよい（ステップ１００１）。

【０１８８】

ＬＵパス定義指示を受信したＭＡは、ＬＵパス定義指示で指定される上位論理デバイスに関する上位論理デバイス管理情報２０３内エントリ３８より、当該上位論理デバイスに対応するデバイスが外部デバイスか否かを判定し（ステップ１００２）、外部デバイスの場合は外部デバイス接続変更処理を行う（ステップ１００３）。

【０１８９】

外部デバイス接続変更処理完了後又はＬＵパス定義指示で指定される上位論理デバイスが外部デバイスに対応付けられていない場合、ＭＡは、当該上位論理デバイスに対してＬＵパス登録を行う。具体的にＭＡは、上位論理デバイス管理情報２０３の当該デバイスエントリのエントリ３５及びエントリ３６にＬＵパス定義指示で指定される対応情報を設定し、ＬＵパス管理情報２０６のＬＵパス定義

指示で指定されるポート１４１に関する空きエントリに、ターゲットＩＤ／ＬＵＮを始めとする構成情報を、ＬＵパス定義指示に従って設定する。登録及び設定が完了したら、ＭＡはその旨を他の各部位へ通知する（ステップ１００４）。

【０１９０】

通知を受信したＰＡは、新たに設定、登録された情報を取込み、通知を受信したＳＴは新たに設定、登録された情報を取り込むと共に、要求元である管理サーバ１１０やストレージ管理者へ完了報告を通知すべく完了報告を出力する（ステップ１００５）。

【０１９１】

図１４は、外部デバイス接続変更処理２５６の処理フローの一例を示す図である。外部デバイスに対応する上位論理デバイスに対してＬＵパス定義される際に、当該外部デバイスに対してストレージ管理者が要求する機能レベルに基づきＣＡ経由での入出力処理が必要であると判断された場合又はＭＡがＳＴを介してストレージ管理者もしくは管理サーバ１１０から上位論理デバイスに対する機能要求レベルの変更指示を受信した場合に、初期状態のＰＡ直結接続から接続形態を切り替えるため、外部デバイス接続変更処理２５６が実行される。

【０１９２】

機能要求レベルの変更指示を受信したＭＡは、当該変更指示内の情報によって特定される上位論理デバイスの接続形態をＣＡ経由とすべきか否かを、当該変更指示に基づいて判断する（ステップ１４０１）。ＣＡ経由とすべきである場合には、ＭＡは、当該上位論理デバイスのアクセスモードがＰＡ直結モードとなっているかどうかを上位論理デバイス管理情報２０３を参照して判断する（ステップ１４０２）。

【０１９３】

該当する上位論理デバイスがＰＡ直結モードである場合には、ＭＡは当該上位論理デバイスの接続形態を「ＰＡ直結」から「ＣＡ経由」へ変更する（ステップ１４０３）。既にＣＡ経由モードとなっている場合には変更は必要ないので処理を終了する。

【０１９４】

ステップ１４０１で当該上位論理デバイスの接続形態をＰＡ直結とすべき場合、ＭＡは上位論理デバイス管理情報２０３を参照して当該上位論理デバイスのデバイスアクセスモードがＣＡ経由となっているかＰＡ直結となっているかを判断する（ステップ１４０４）。

【０１９５】

当該上位論理デバイスのデバスアクセスモードがＣＡ経由となっている場合、ＭＡは接続形態を「ＣＡ経由」から「ＰＡ直結」へ変更する必要がある。そこでまずＭＡは、当該上位論理デバイスの上位論理デバイス管理情報２０３内のエントリ３７を「ＰＡ直結移行中」へ変更し（ステップ１４０５）、当該上位論理デバイスに対応付けられている下位論理デバイスを管理しているＣＡを上位論理デバイス管理情報２０３を参照して特定する。そして特定されたＣＡに当該上位論理デバイスの接続形態がＣＡ経由からＰＡ直結に変更される旨を通知する（ステップ１４０６）。

【０１９６】

通知を受けたＣＡは、対応する下位論理デバイスの全データについてディスクキャッシュ１５４を検索しディスクキャッシュをミス化、すなわち、物理デバイス又は外部デバイスへ未更新のデータをディスクキャッシュ１５４から対応デバイスへ書き出してデバイス内のデータを更新した後、更新済みのデータについては即、当該データに割り当てられたキャッシュ領域を解放する。ＣＡは、ディスクキャッシュ１５４の検索をキャッシュ管理情報を用いて、当該下位論理デバイスの先頭から順次行う。又、ＣＡは、検索およびミス化処理が終わった部分については、上位論理デバイス及び下位論理デバイス管理情報の切替進捗ポインタを進めることで進捗を管理する（ステップ１４０７）。

【０１９７】

ディスクキャッシュ１５４の全領域についてミス化処理が完了したら、ＣＡはその旨をＭＡへ報告し（ステップ１４０８）、ＭＡで上位論理デバイス管理情報２０３内のエントリ３７を「ＰＡ直結」へ変更して接続形態変更処理を完了する。

【０１９８】

以上の実施形態によれば、ホスト１００がコマンドを送付するアドレスが付与されるI/F部分（ＰＡ）とファイルアクセス処理を行うNASサーバ部分（ＡＡ）とデバイスの処理を行うキャッシュ制御部（ＣＡ）と外部ストレージ１８０と接続されるI/F部分（ＰＡ）を分離し、相互結合網１７０で接続することで、処理負荷を分散可能なNASシステムが実現できる。

【０１９９】

また、本実施形態によれば、ホスト１００がコマンドを送付するアドレスが付与されるI/F部分（ＰＡ）とファイルアクセス処理を行うNASサーバ部分（ＡＡ）とデバイスの処理を行うキャッシュ制御部（ＣＡ）と外部ストレージ１８０と接続されるI/F部分（ＰＡ）とで「連携」モードを設け、大量のデータがホスト１００とＡＡで送受されるような場合、対ホストのI/F部分（ＰＡ）とキャッシュ制御部（ＣＡ）もしくは外部ストレージ１８０と接続されるI/F部分（ＰＡ）間で直接データを送受することができ、ＡＡの処理負荷を軽減できると同時に、ホストに対するコマンド応答が早いNASシステムが実現できる。

【０２００】

また、本実施形態によれば、NAS管理情報２２３で、ＡＡが管理するファイルとそれにアクセスするためにホスト１００がコマンドを送信する先であるI/F部分のIPアドレスの対応関係を、１つのＰＡに対し複数のＡＡが処理するようにエントリ１６０２に設定でき、ホスト１００からは1台のNASシステムにアクセスするように見えながら、複数のＡＡで処理を実施することで処理負荷を軽減し、ホスト１００に対するコマンド応答が早いNASシステムが実現できる。

【０２０１】

また、本実施形態によれば、NAS管理情報２２３でＡＡが管理するファイルとそれにアクセスするためにホスト１００がコマンドを送信する先であるI/F部分のIPアドレスの対応関係を、１つのＡＡが複数のIPアドレスを持っているようにエントリ１６０２に設定でき、ホスト１００からは複数のNASシステムが存在するように見えながら、複数のＰＡまたはポート１４１でI/F処理を実施することで、処理負荷を軽減し、ホスト１００に対するコマンド応答が早いNASシステムが実現できる。

【０２０２】

また、本実施形態によれば、NAS管理情報２２３でＡＡが管理するファイルとそれにアクセスするためにホスト１００がコマンドを送信する先であるI/F部分のIPアドレスの対応関係を、１つのＡＡが複数のIPアドレスを持っているようにエントリ１６０２に設定でき、ホスト１００からは複数のNASシステムが存在するように見え、NASシステム毎に用途を使い分けることができるNASシステムが実現できる。

【０２０３】

また、本実施形態によれば、NAS管理情報２２３でＡＡが管理するファイルとそれにアクセスするためにホスト１００がコマンドを送信する先であるI/F部分のIPアドレスの対応関係の一部を他のＡＡに移すことができ、これにより、ホスト１００に影響を及ぼすことなく、処理負荷が大きくなってきたＡＡを増設し処理負荷を軽減し、ホストに対するコマンド応答を向上できるNASシステムが実現できる。

【０２０４】

また、本実施例によれば、NAS管理情報２２３でＡＡが管理するファイルとそれにアクセスするためにホスト１００がコマンドを送信する先であるI/F部分のIPアドレスの対応関係の一部を変更することができ、これにより、処理負荷が大きくなってきたＰＡを増設もしくは処理負荷が少ないＰＡへ負荷を分散させ、これにより処理負荷を軽減し、ホストに対するコマンド応答を向上できるNASシステムが実現できる。

【０２０５】

また本実施形態によれば、NAS管理情報２２３のＡＡに関するエントリ全てを新規のＡＡのエントリにコピーすることで、ホスト１００に影響を及ぼすことなく、ＡＡを処理能力の高い新しいパッケージに交換でき、処理負荷を軽減し、ホストに対するコマンド応答を向上できるNASシステムが実現できる。

【０２０６】

また本実施形態によれば、ポート管理情報２２４のポートに付与されているIPアドレスの一部を別のポート１４１のエントリに移動することで、ホスト１００

の設定を変更することなく、ポート１４１へのアクセスを分散させ、処理負荷を軽減し、ホストに対するコマンド応答を向上できるNASシステムが実現できる。

【０２０７】

また本実施形態によれば、ＡＡに障害が発生した場合に、NAS管理情報２２３に予め設定してあった代替ＡＡへ当該ＡＡの管理情報を引き継ぐことができ、ホスト１００の設定を変更することなく処理を継続可能なNASシステムが実現できる。

【０２０８】

また本実施形態によれば、ＰＡのポート１４１に障害が発生した場合に、ポート管理情報２２４の当該ポート１４１に関する情報を全てを使用されていないポート１４１のエントリにコピーするか、使用されているポート１４１に当該ポートに関する情報の一部を移動することで、ホスト１００の設定を変更することなく使用できるNASシステムが実現できる。

【０２０９】

また、本実施形態によれば、ＣＡが外部ストレージを自デバイスとしてＡＡに提供できるので、ストレージ１３０にディスク装置１５７を搭載せずに、安い外部ストレージを利用することができ、前述の本実施形態の効果を持った低価格のNASシステムを実現することができる。

【０２１０】

また、本実施形態によれば、ＣＡを介さずに、ＡＡが直接外部ストレージを使用することができ、前述の本実施形態の効果を持った低価格のNASシステムを実現することができる。

【０２１１】

【発明の効果】

本発明によれば、NASの使用負荷が上がった場合にも、レスポンスが低下しないように、処理負荷が高い部分の処理を分散するよう制御可能である。

【０２１２】

また、本発明によれば、障害が発生した場合にも障害発生部分の情報を他に引き継ぐことができ、計算機から当該NASへのアクセスを継続もしくは計算機の設

定を変更しないで継続利用できる。

【図面の簡単な説明】

【図１】

システムのハードウェア構成の一例を示す図である。

【図２】

ストレージ１３０のソフトウェア構成の一例を示す図である。

【図３】

上位論理デバイス管理情報の一例を示す図である。

【図４】

ＬＵパス管理情報の一例を示す図である。

【図５】

下位論理デバイス管理情報の一例を示す図である。

【図６】

物理デバイス管理情報の一例を示す図である。

【図７】

外部デバイス管理情報の一例を示す図である。

【図８】

外部デバイス定義処理の一例を示す図である。

【図９】

論理デバイス定義処理の一例を示す図である。

【図１０】

ＬＵパス定義処理の一例を示す図である。

【図１１】

要求振り分け処理の一例を示す図である。

【図１２】

コマンド処理の一例を示す図である。

【図１３】

ライトアフタ処理の一例を示す図である。

【図１４】

外部デバイス接続変更処理の一例を示す図である。

【図１５】

管理端末のソフトウェア構成の一例を示す図である。

【図１６】

NAS管理情報の一例を示す図である。

【図１７】

ポート管理情報の一例を示す図である。

【図１８】

NAS管理情報変更処理の一例を示す図である。

【図１９】

ポート管理情報変更処理の一例を示す図である。

【図２０】

NAS／ポート障害処理の一例を示す図である。

【図２１】

ファイルアクセス処理の一例を示す図である。

【図２２】

連携モード処理の一例を示す図である。

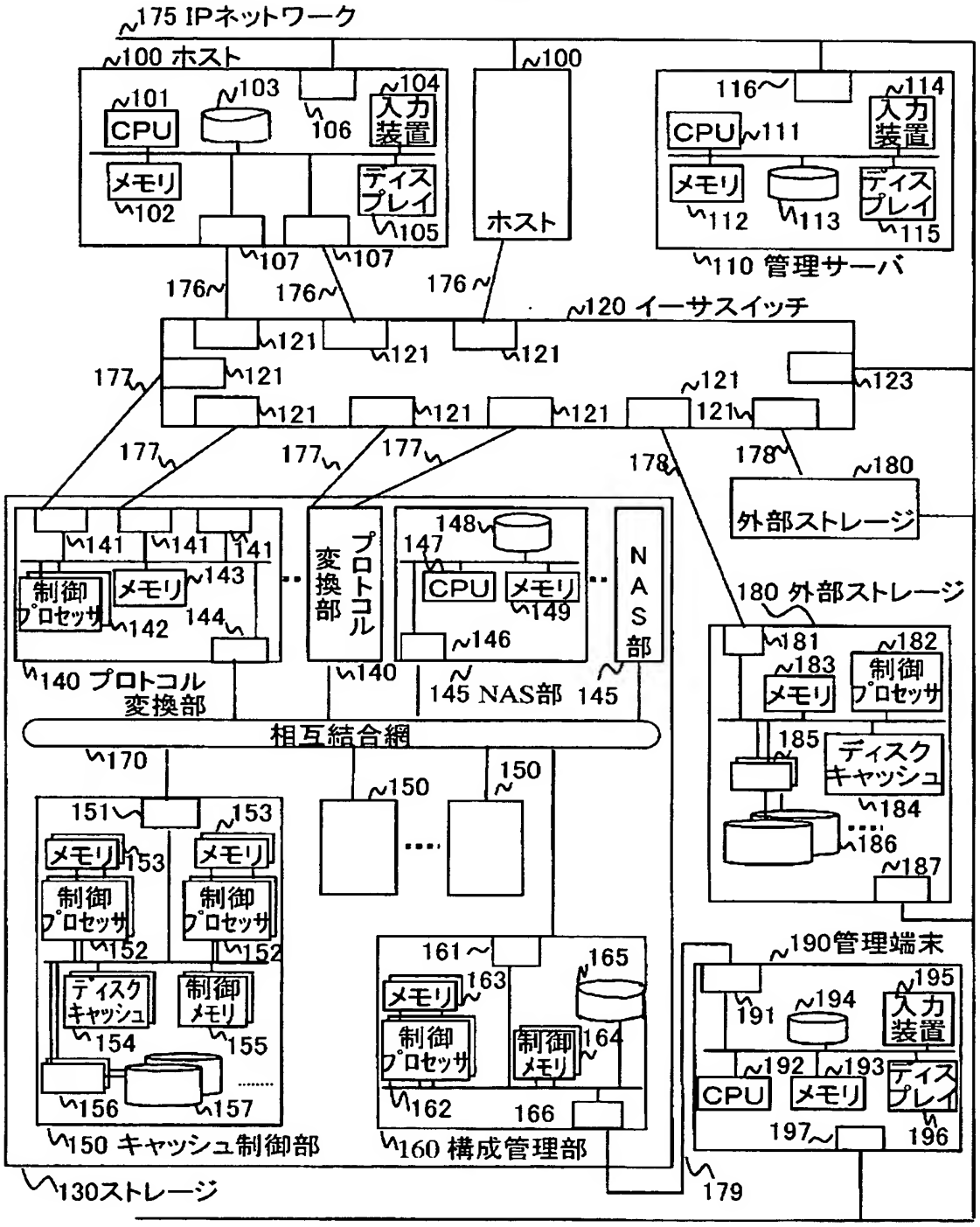
【符号の説明】

１００…ホスト、１１０…管理サーバ、１２０…スイッチ、１３０…ストレージ装置、１４０…プロトコル変換部、１４５…NAS部、１５０…キャッシュ制御部、１６０…構成管理部、１７０…相互結合網、１８０…外部ストレージ装置、１９０…管理端末。

【書類名】　図面

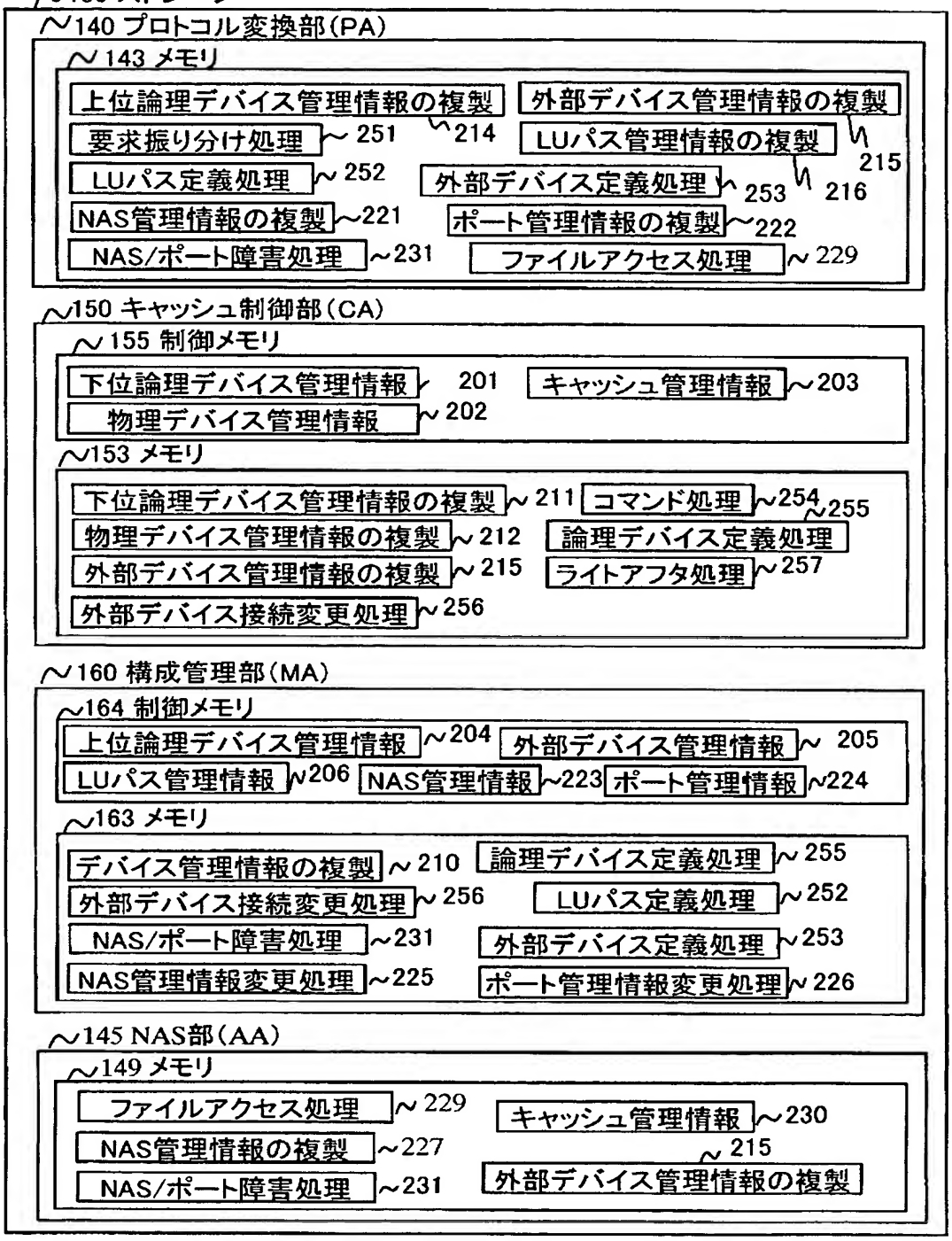
【図１】

図1


175 IPネットワーク
100 ホスト
101
CPU
102
メモリ
103
106
104
入力装置
105
ディスプレイ
107
107
100
ホスト
116
CPU
111
メモリ
112
113
114
入力装置
115
ディスプレイ
110 管理サーバ
176
176
176
120 イーサスイッチ
121
121
121
121
121
121
121
121
121
123
177
177
177
177
178
178
180
外部ストレージ
141
141
141
制御プロセッサ
142
メモリ
143
144
140 プロトコル変換部
プロトコル変換部
140
148
147
CPU
メモリ
149
146
145 NAS部
NAS部
145
180 外部ストレージ
181
182
183
メモリ
制御プロセッサ
185
ディスクキャッシュ
184
186
187
相互結合網
170
150
150
151
153
メモリ
153
メモリ
制御プロセッサ
制御プロセッサ
152
152
ディスクキャッシュ
154
制御メモリ
155
156
157
150 キャッシュ制御部
161
165
メモリ
163
制御プロセッサ
制御メモリ
164
162
166
160 構成管理部
190管理端末
191
194
195
入力装置
192
193
CPU
メモリ
ディスプレイ
196
197
179
130ストレージ

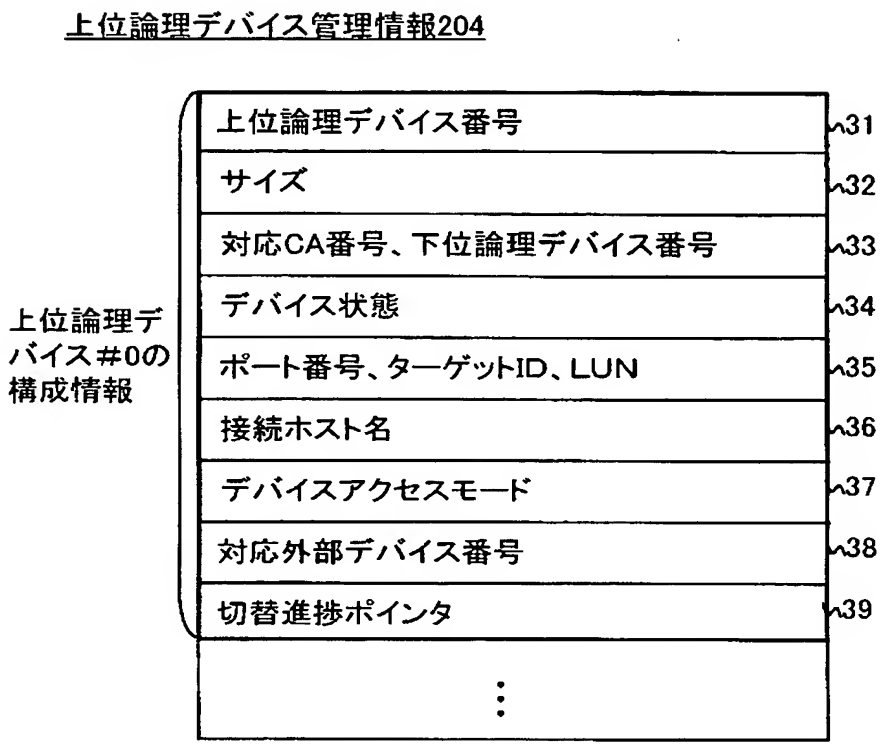
【図２】

図2


130 ストレージ
140 プロトコル変換部(PA)
143 メモリ
上位論理デバイス管理情報の複製
外部デバイス管理情報の複製
要求振り分け処理 251
214
LUパス管理情報の複製
215
LUパス定義処理 252
外部デバイス定義処理 253
216
NAS管理情報の複製 221
ポート管理情報の複製 222
NAS/ポート障害処理 231
ファイルアクセス処理 229
150 キャッシュ制御部(CA)
155 制御メモリ
下位論理デバイス管理情報 201
キャッシュ管理情報 203
物理デバイス管理情報 202
153 メモリ
下位論理デバイス管理情報の複製 211
コマンド処理 254
255
物理デバイス管理情報の複製 212
論理デバイス定義処理
外部デバイス管理情報の複製 215
ライトアフタ処理 257
外部デバイス接続変更処理 256
160 構成管理部(MA)
164 制御メモリ
上位論理デバイス管理情報 204
外部デバイス管理情報 205
LUパス管理情報 206
NAS管理情報 223
ポート管理情報 224
163 メモリ
デバイス管理情報の複製 210
論理デバイス定義処理 255
外部デバイス接続変更処理 256
LUパス定義処理 252
NAS/ポート障害処理 231
外部デバイス定義処理 253
NAS管理情報変更処理 225
ポート管理情報変更処理 226
145 NAS部(AA)
149 メモリ
ファイルアクセス処理 229
キャッシュ管理情報 230
NAS管理情報の複製 227
215
NAS/ポート障害処理 231
外部デバイス管理情報の複製

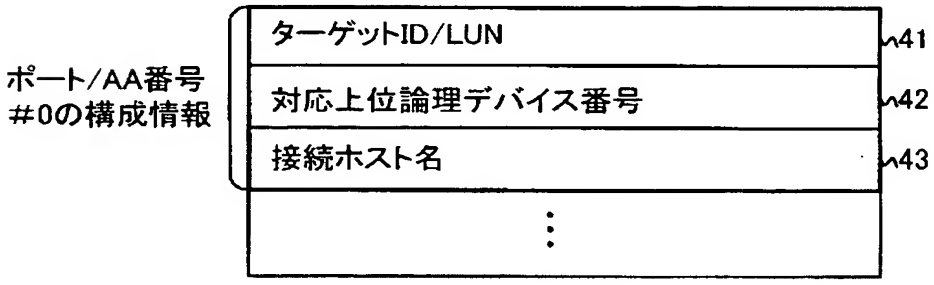
【図３】

図3


上位論理デバイス管理情報204
上位論理デバイス番号
サイズ
対応CA番号、下位論理デバイス番号
デバイス状態
ポート番号、ターゲットID、LUN
接続ホスト名
デバイスアクセスモード
対応外部デバイス番号
切替進捗ポインタ
上位論理デバイス#0の構成情報
31
32
33
34
35
36
37
38
39

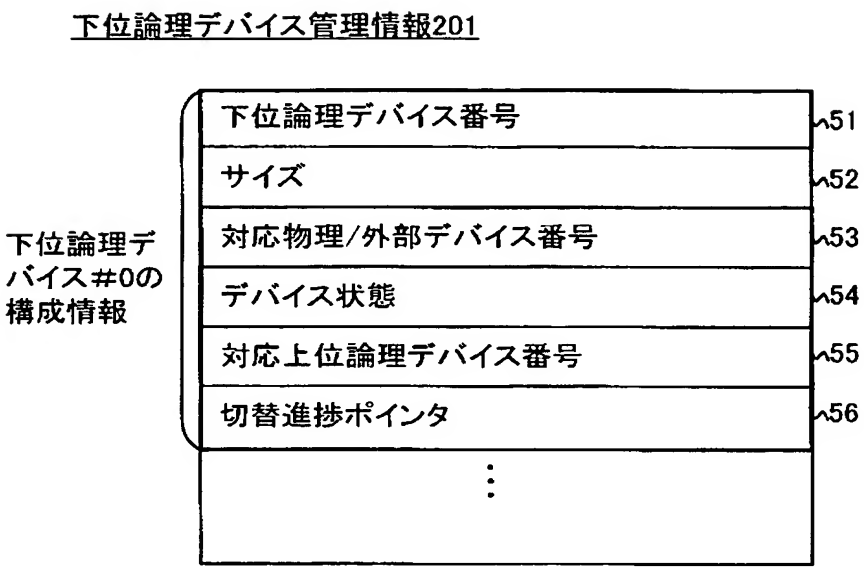
【図４】

図4


LUパス管理情報206
ターゲットID/LUN
対応上位論理デバイス番号
接続ホスト名
ポート/AA番号#0の構成情報
41
42
43

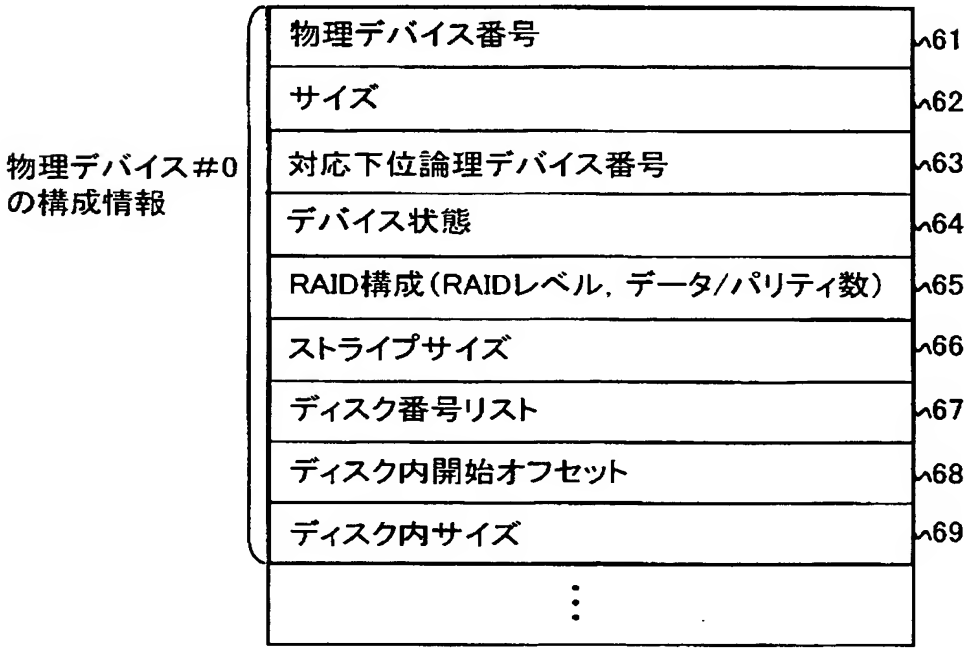
【図５】

図5


下位論理デバイス管理情報201
下位論理デバイス#0の構成情報
下位論理デバイス番号
51
サイズ
52
対応物理/外部デバイス番号
53
デバイス状態
54
対応上位論理デバイス番号
55
切替進捗ポインタ
56

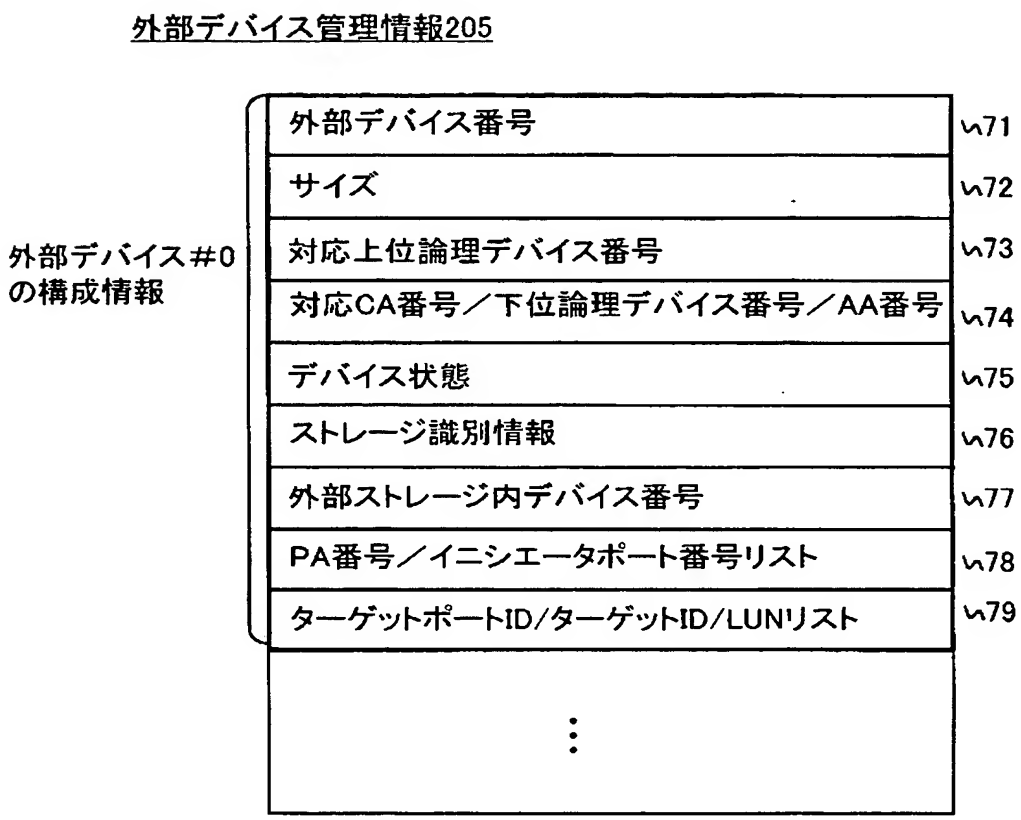
【図６】

図6


物理デバイス管理情報202
物理デバイス#0の構成情報
物理デバイス番号
61
サイズ
62
対応下位論理デバイス番号
63
デバイス状態
64
RAID構成（RAIDレベル，データ/パリティ数）
65
ストライプサイズ
66
ディスク番号リスト
67
ディスク内開始オフセット
68
ディスク内サイズ
69

【図７】

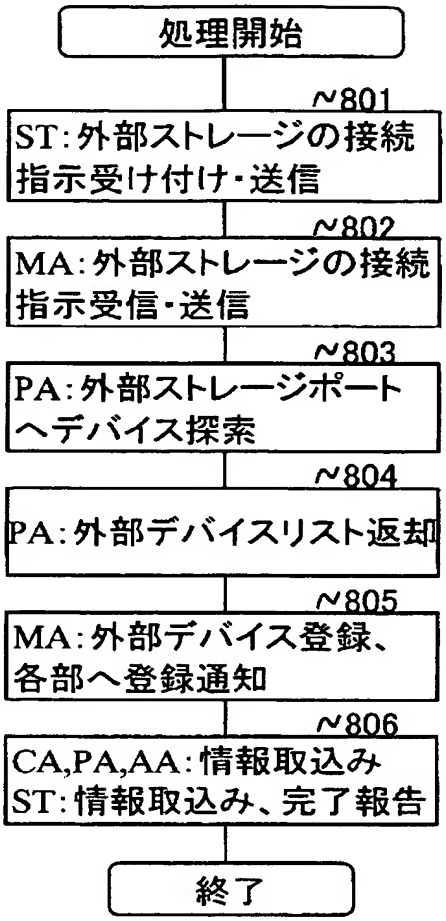
図7


外部デバイス管理情報205
外部デバイス#0の構成情報
外部デバイス番号
サイズ
対応上位論理デバイス番号
対応CA番号／下位論理デバイス番号／AA番号
デバイス状態
ストレージ識別情報
外部ストレージ内デバイス番号
PA番号／イニシエータポート番号リスト
ターゲットポートID/ターゲットID/LUNリスト
71
72
73
74
75
76
77
78
79

【図８】

図8

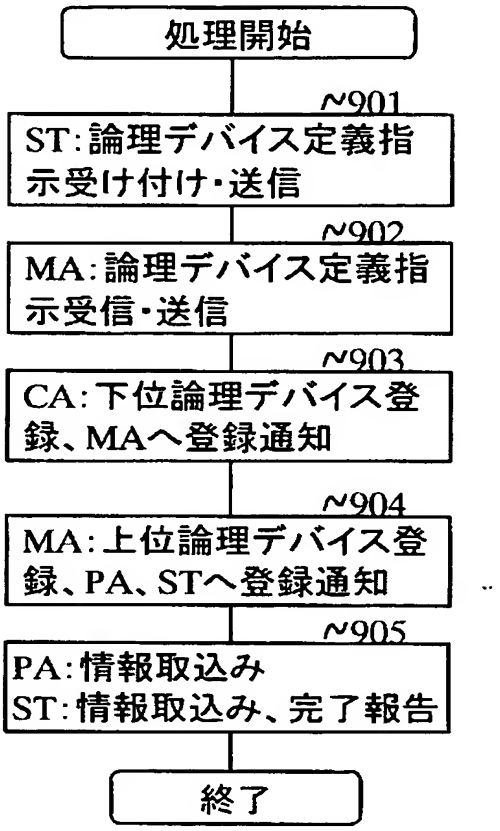
外部デバイス定義処理253



【図９】

図９

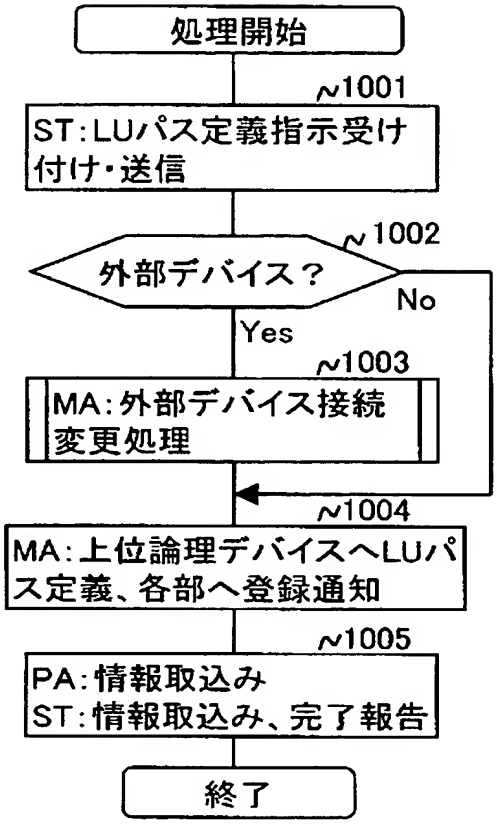
論理デバイス定義処理255



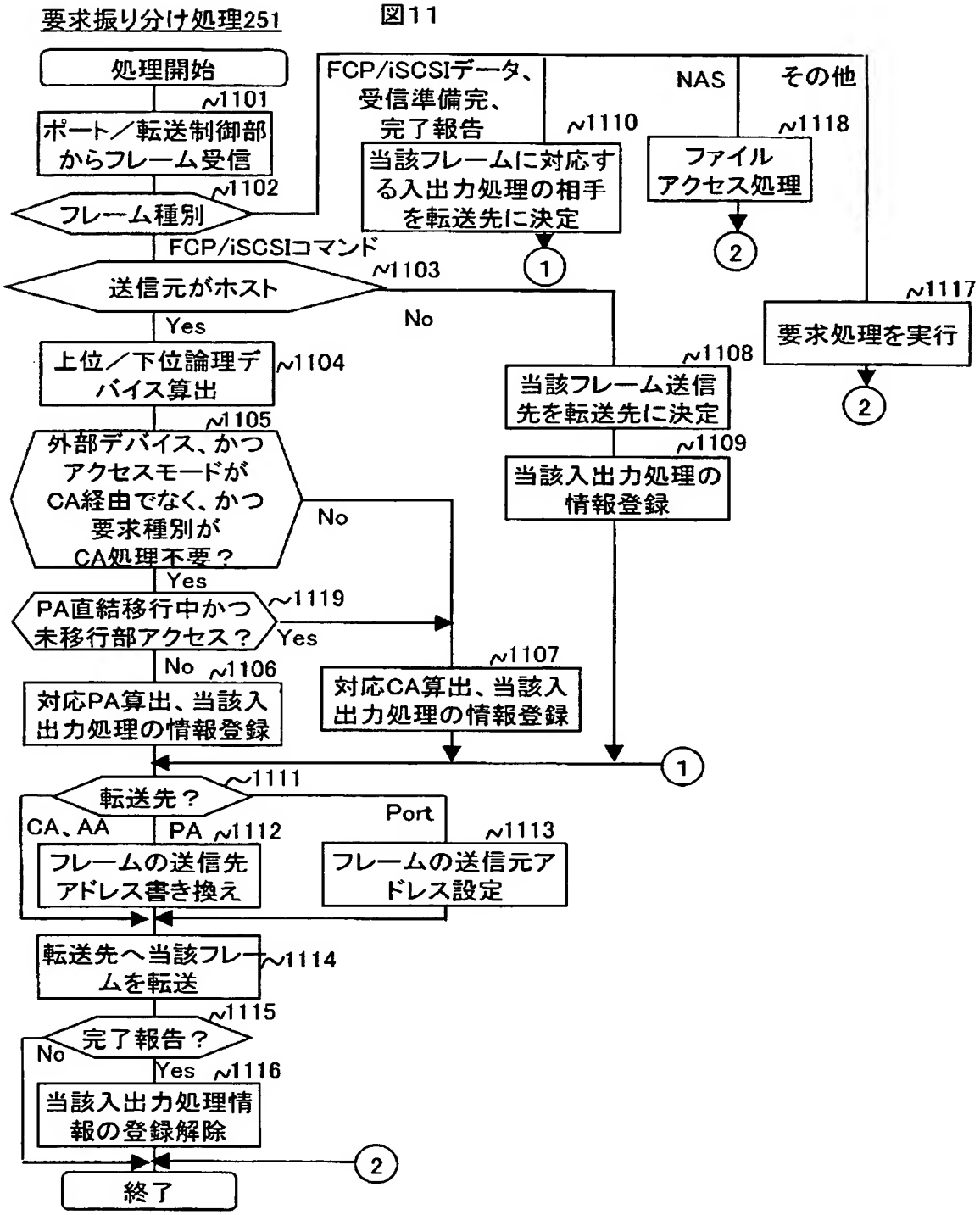
【図１０】

図10

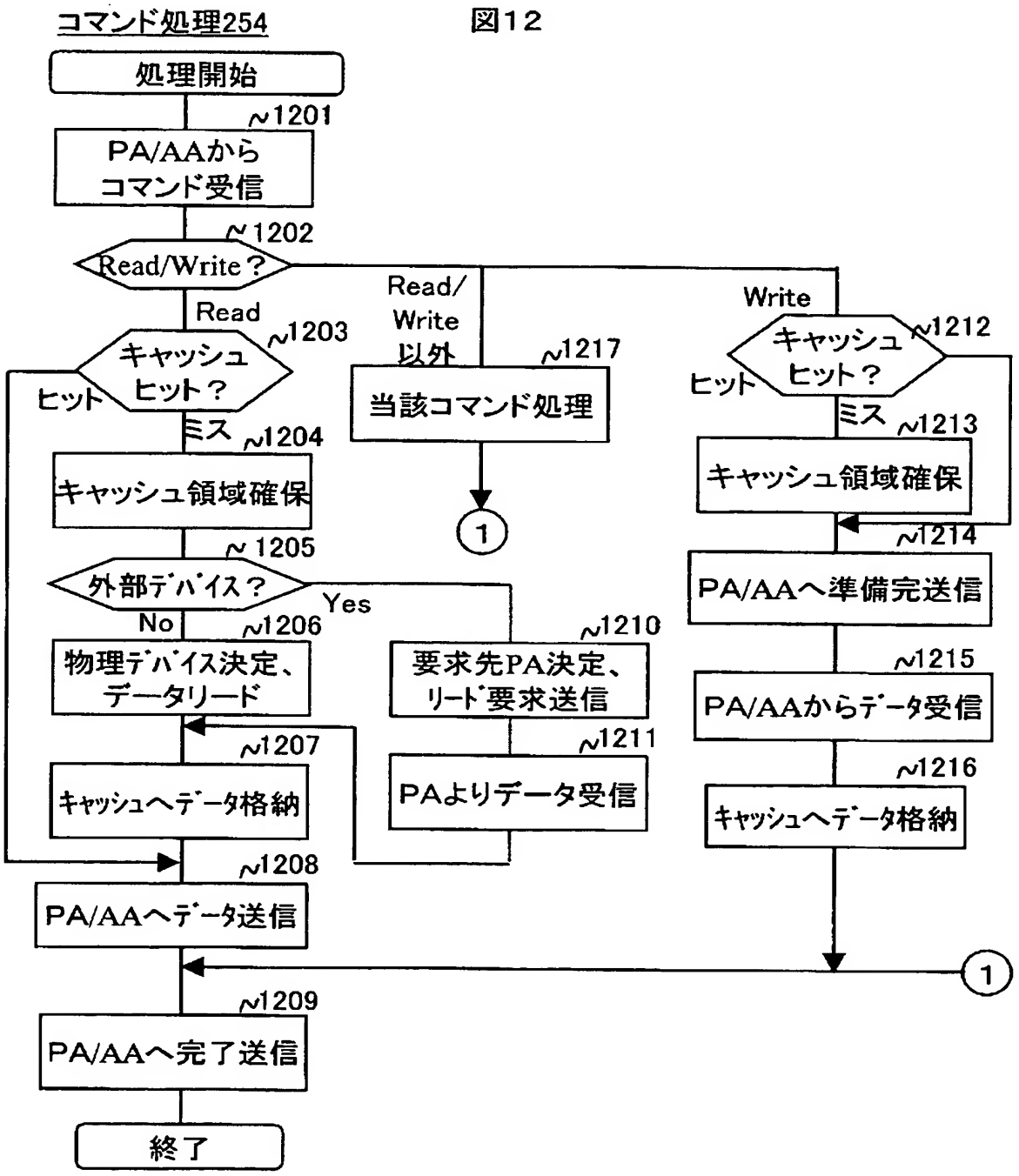
LUパス定義処理252



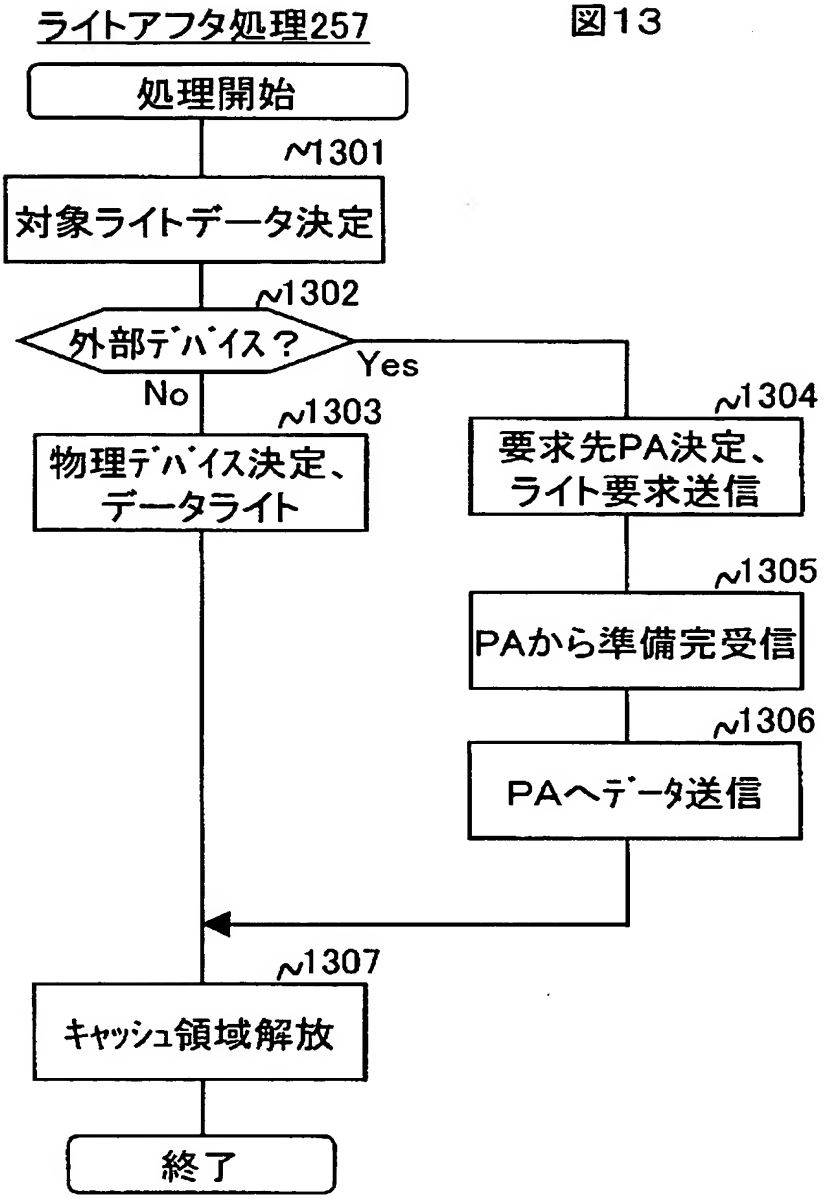
【図１１】


要求振り分け処理251
図11
処理開始
~1101
ポート／転送制御部からフレーム受信
~1102
フレーム種別
FCP/iSCSIデータ、受信準備完、完了報告
NAS
その他
~1110
当該フレームに対応する入出力処理の相手を転送先に決定
~1118
ファイルアクセス処理
FCP/iSCSIコマンド
~1103
送信元がホスト
No
Yes
~1117
要求処理を実行
~1104
上位／下位論理デバイス算出
~1108
当該フレーム送信先を転送先に決定
~1105
外部デバイス、かつアクセスモードがCA経由でなく、かつ要求種別がCA処理不要？
~1109
当該入出力処理の情報登録
No
Yes
~1119
PA直結移行中かつ未移行部アクセス？
Yes
No
~1106
対応PA算出、当該入出力処理の情報登録
~1107
対応CA算出、当該入出力処理の情報登録
~1111
転送先？
CA、AA
PA
~1112
フレームの送信先アドレス書き換え
Port
~1113
フレームの送信元アドレス設定
~1114
転送先へ当該フレームを転送
~1115
完了報告？
No
Yes
~1116
当該入出力処理情報の登録解除
終了
1
2

【図１２】


コマンド処理254
図12
処理開始
1201
PA/AAから
コマンド受信
1202
Read/Write？
Read
Read/
Write
以外
Write
1203
キャッシュ
ヒット？
ヒット
ミス
1204
キャッシュ領域確保
1205
外部ﾃﾞﾊﾞｲｽ？
Yes
No
1206
物理ﾃﾞﾊﾞｲｽ決定、
ﾃﾞｰﾀﾘｰﾄﾞ
1207
ｷｬｯｼｭへﾃﾞｰﾀ格納
1208
PA/AAへﾃﾞｰﾀ送信
1209
PA/AAへ完了送信
終了
1217
当該コマンド処理
1
1210
要求先PA決定、
ﾘｰﾄﾞ要求送信
1211
PAよりデータ受信
1212
キャッシュ
ヒット？
ヒット
ミス
1213
キャッシュ領域確保
1214
PA/AAへ準備完送信
1215
PA/AAからﾃﾞｰﾀ受信
1216
ｷｬｯｼｭへﾃﾞｰﾀ格納
1

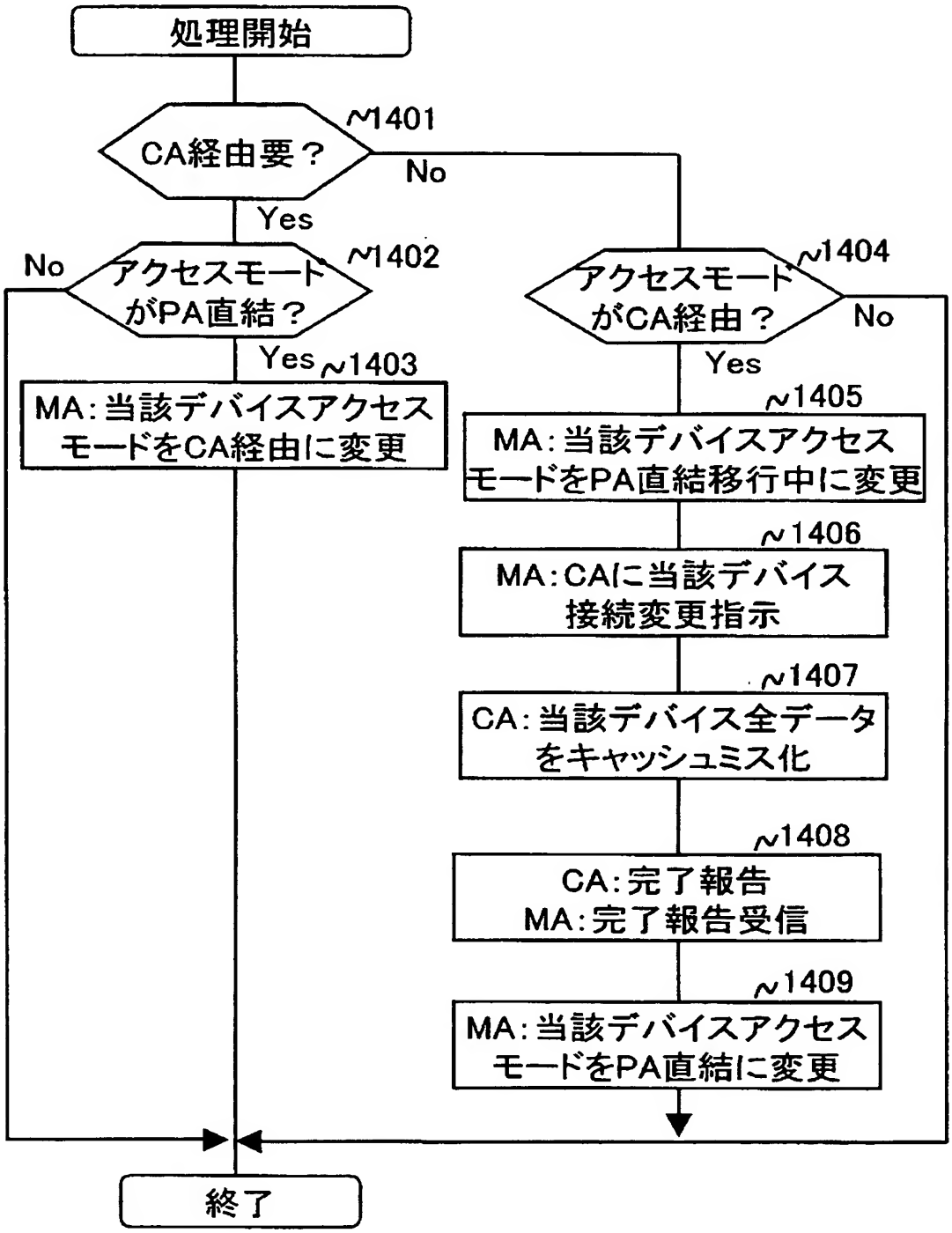
【図１３】


ライトアフタ処理257
図13
処理開始
1301
対象ライトデータ決定
1302
外部ﾃﾞﾊﾞｲｽ？
Yes
No
1303
物理ﾃﾞﾊﾞｲｽ決定、
データライト
1304
要求先PA決定、
ライト要求送信
1305
PAから準備完受信
1306
PAへﾃﾞｰﾀ送信
1307
ｷｬｯｼｭ領域解放
終了

【図１４】

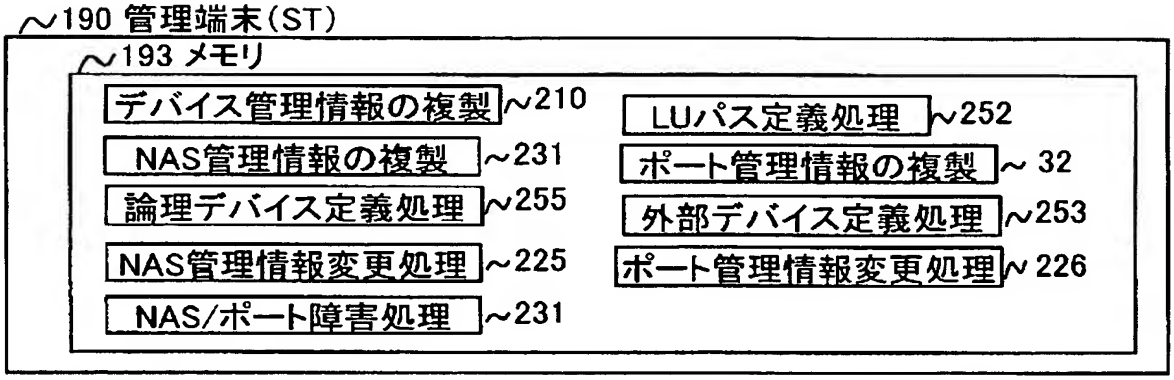
図14

外部デバイス接続変更処理256



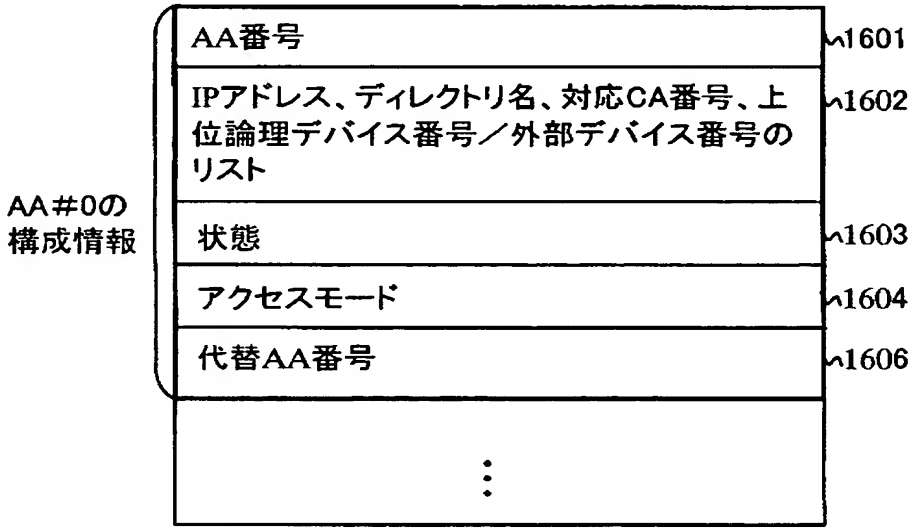
【図１５】

図15



【図１６】

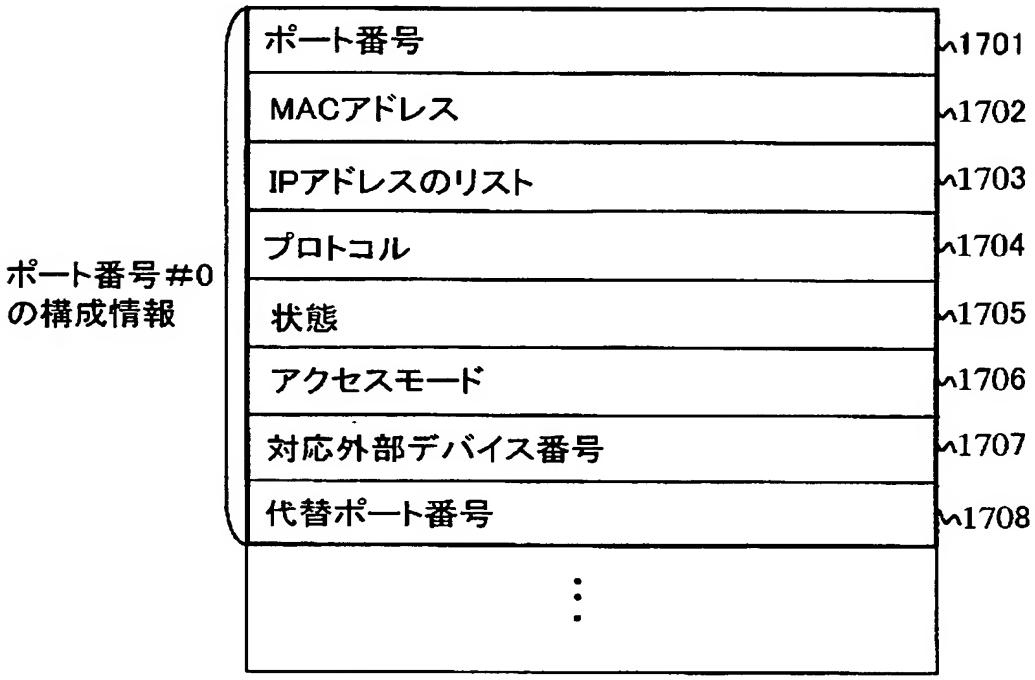
図16



【図１７】

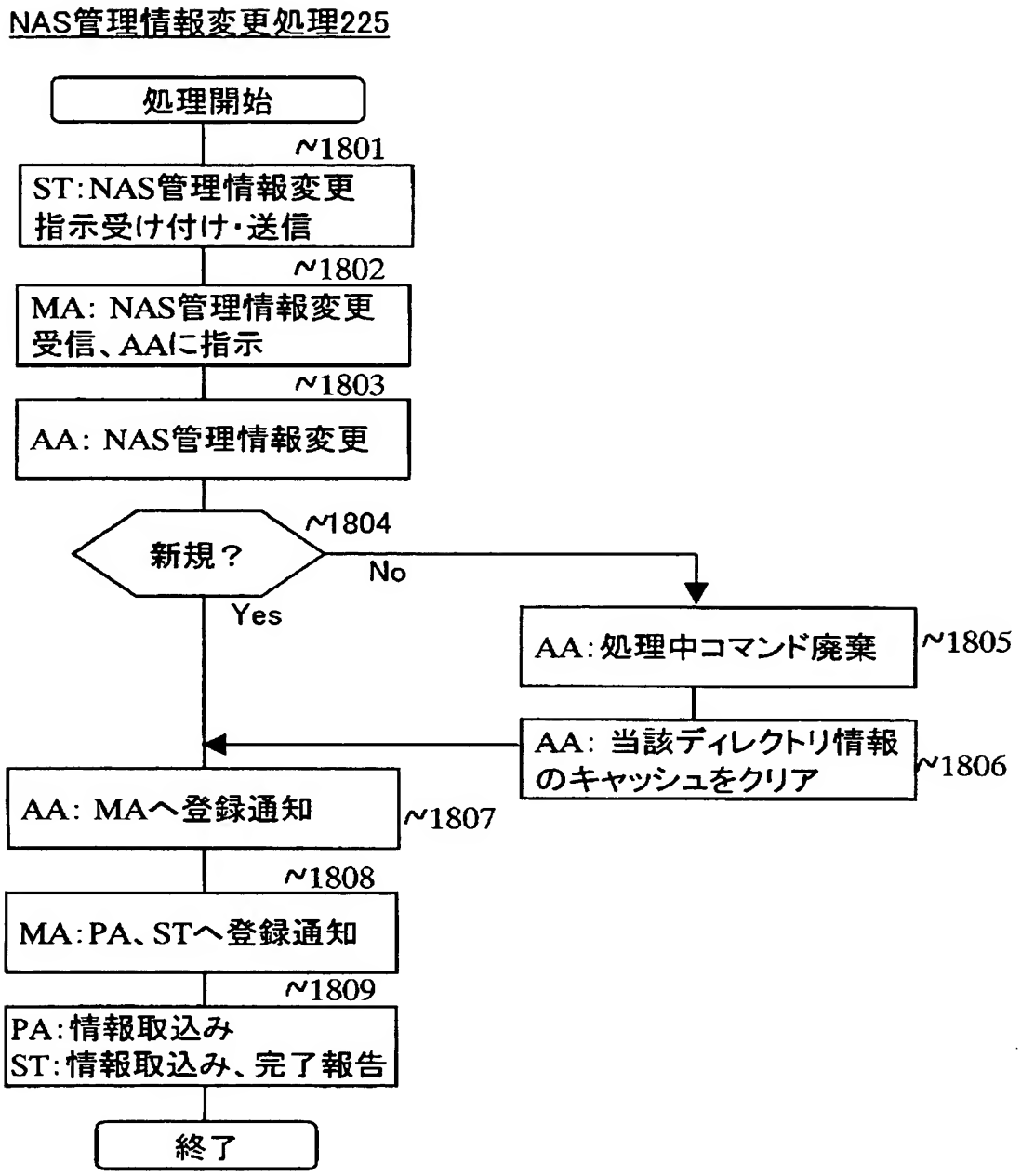
図17

ポート管理情報 224



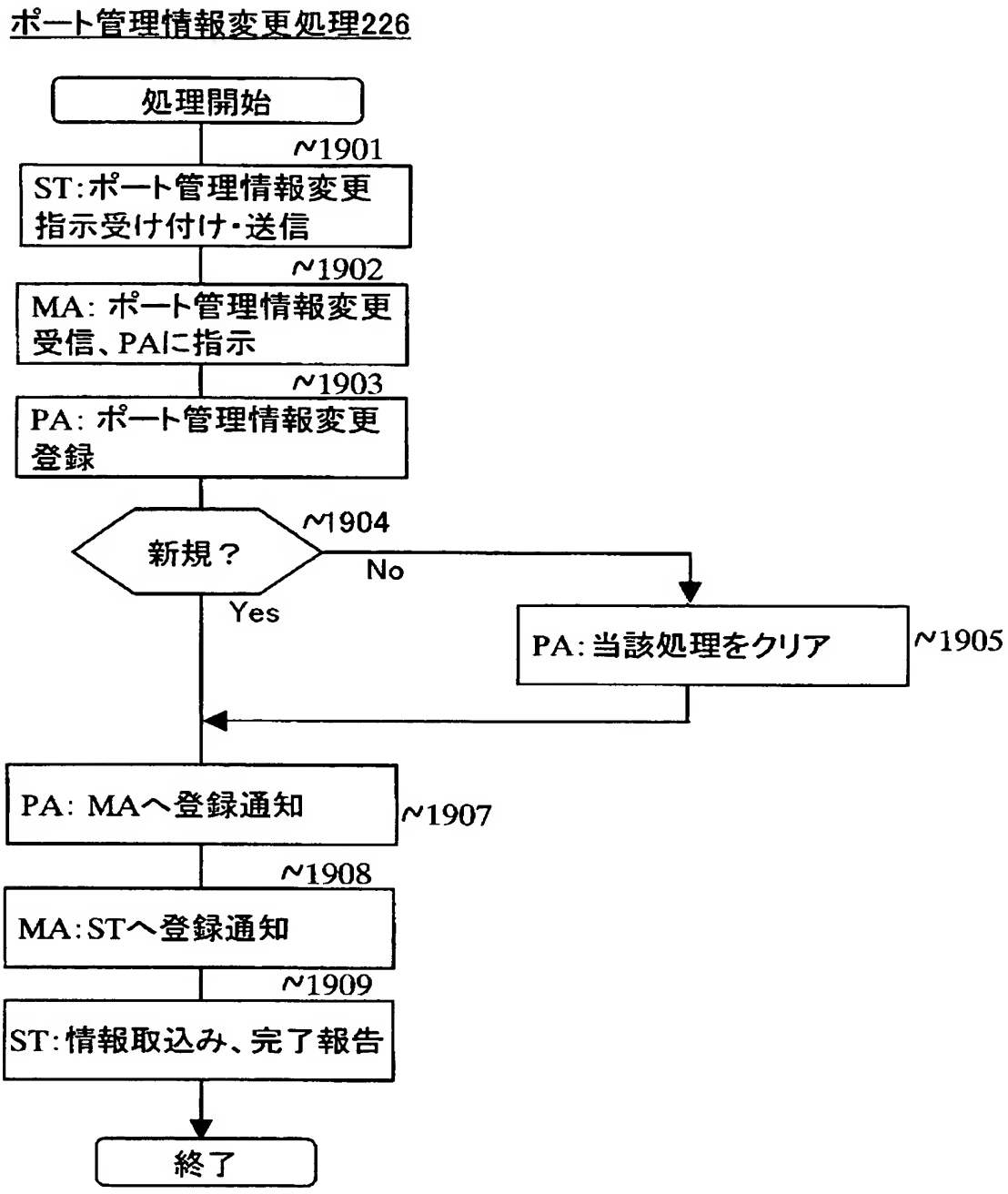
【図１８】

図18


NAS管理情報変更処理225
処理開始
~1801
ST:NAS管理情報変更
指示受け付け・送信
~1802
MA: NAS管理情報変更
受信、AAに指示
~1803
AA: NAS管理情報変更
~1804
新規？
No
Yes
AA:処理中コマンド廃棄
~1805
AA: 当該ディレクトリ情報
のキャッシュをクリア
~1806
AA: MAへ登録通知
~1807
~1808
MA:PA、STへ登録通知
~1809
PA:情報取込み
ST:情報取込み、完了報告
終了

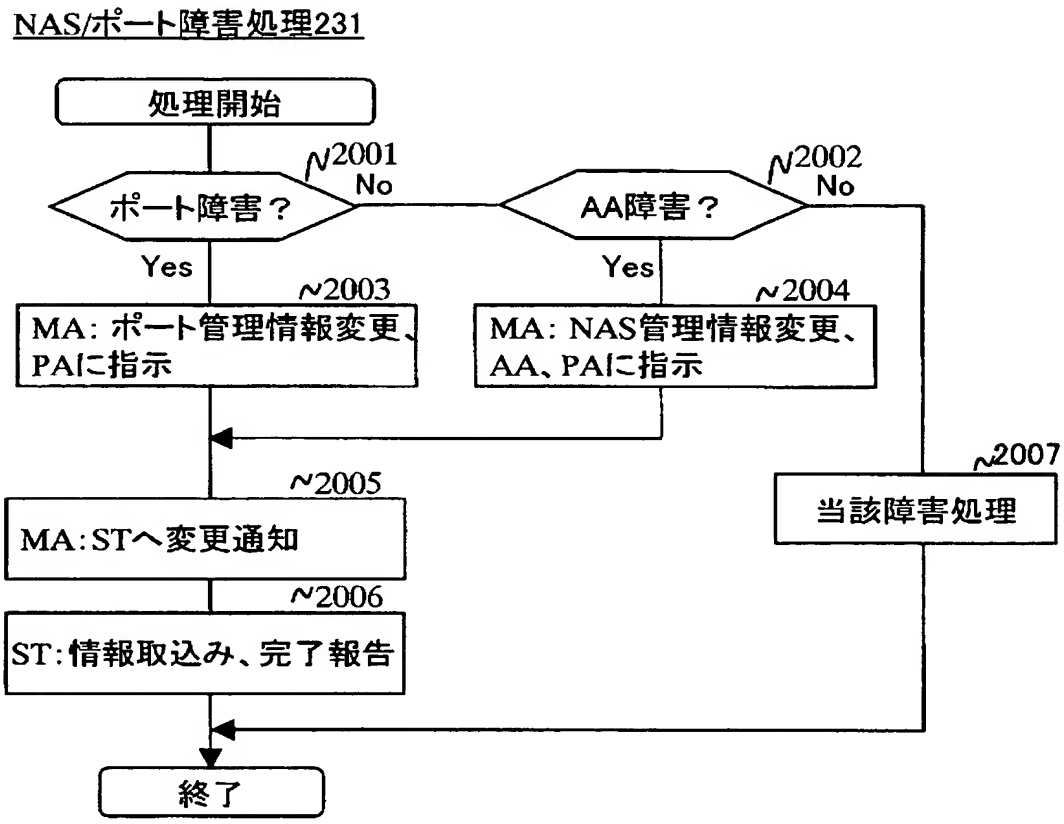
【図１９】

図19

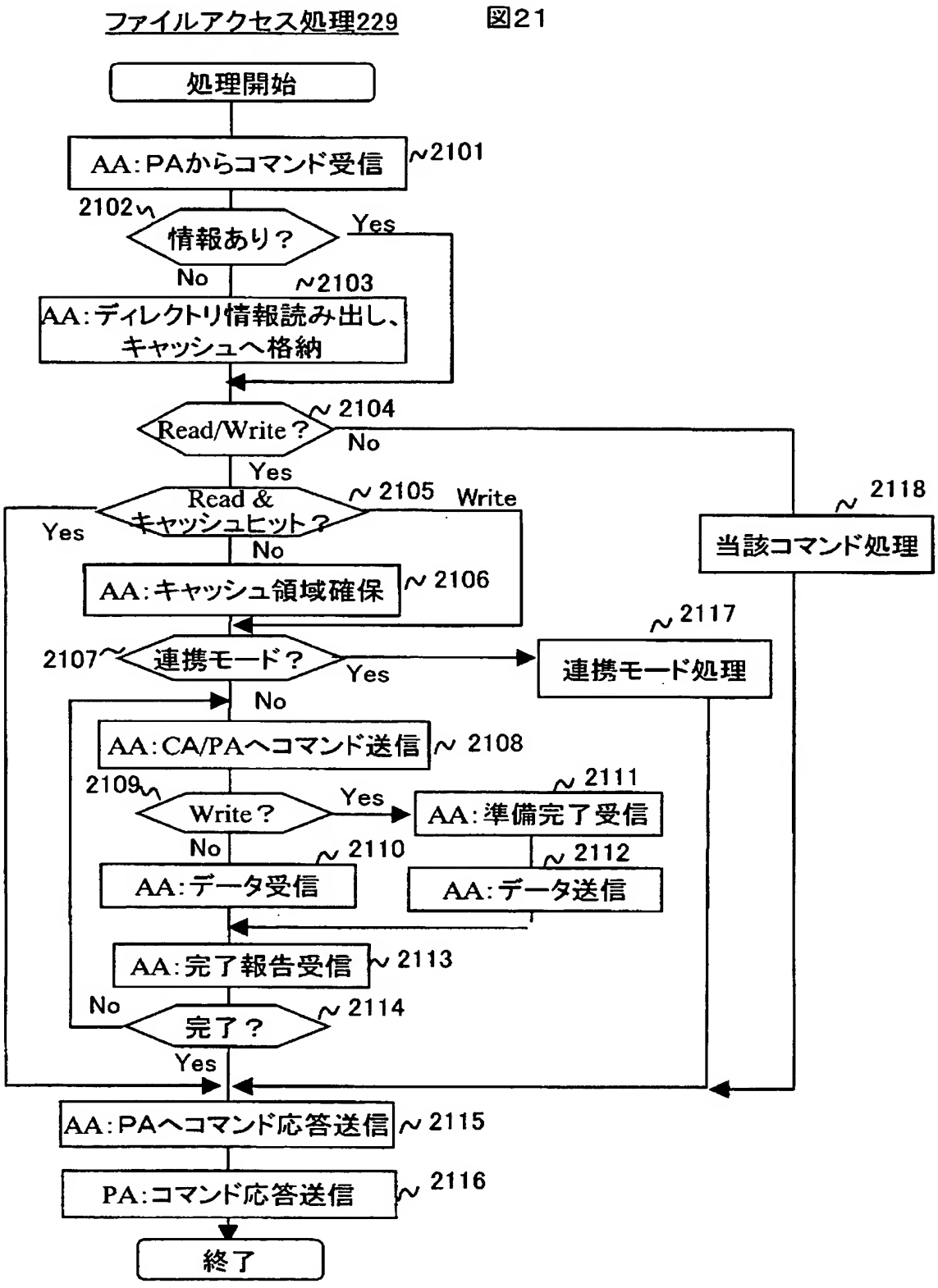

ポート管理情報変更処理226
処理開始
~1901
ST:ポート管理情報変更指示受け付け・送信
~1902
MA: ポート管理情報変更受信、PAに指示
~1903
PA: ポート管理情報変更登録
~1904
新規？
No
Yes
PA:当該処理をクリア
~1905
PA: MAへ登録通知
~1907
~1908
MA:STへ登録通知
~1909
ST:情報取込み、完了報告
終了

【図２０】

図20

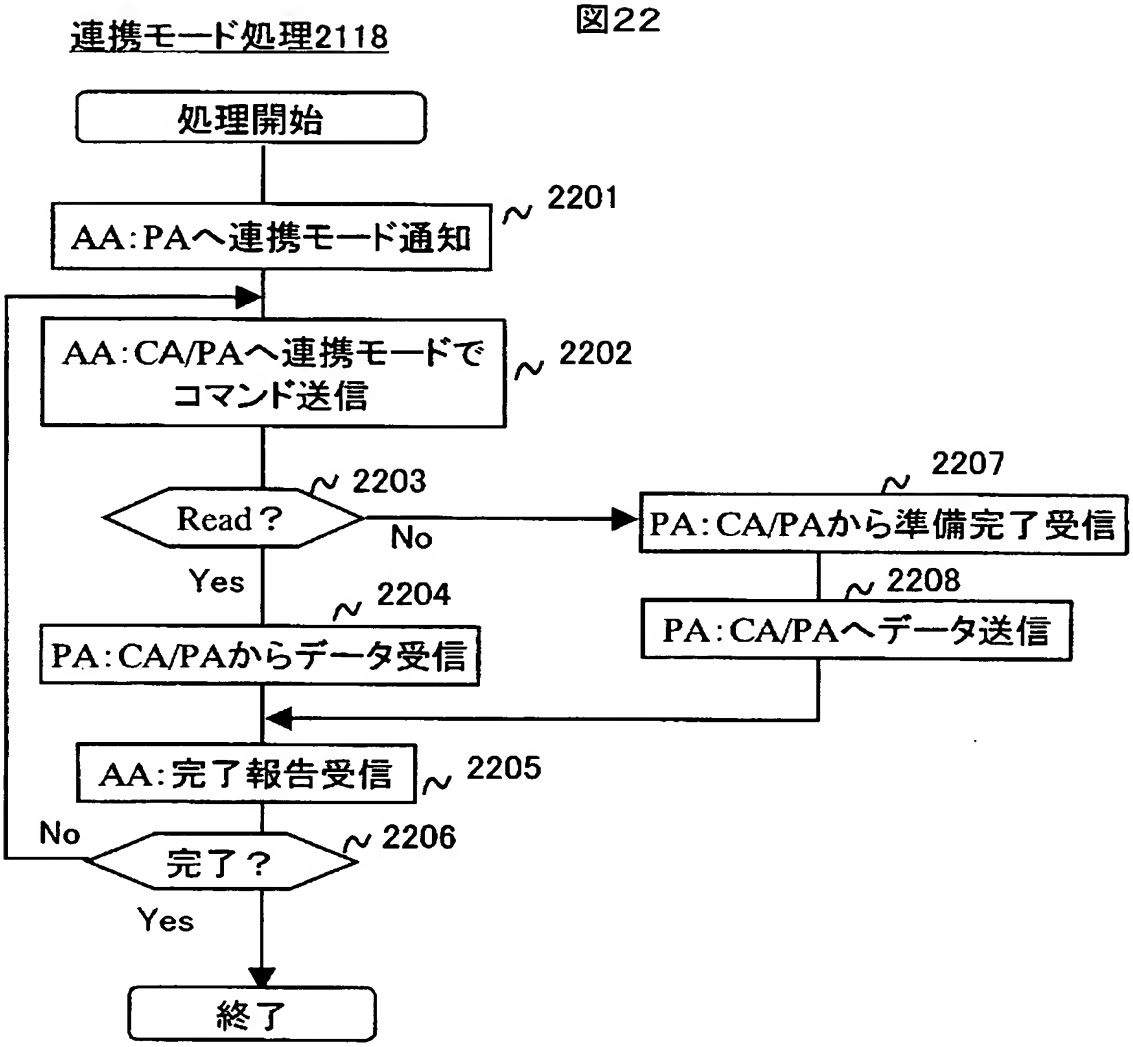

NAS/ポート障害処理231
処理開始
~2001
ポート障害？
No
Yes
~2002
AA障害？
No
Yes
~2003
MA: ポート管理情報変更、PAに指示
~2004
MA: NAS管理情報変更、AA、PAに指示
~2005
MA:STへ変更通知
~2006
ST:情報取込み、完了報告
~2007
当該障害処理
終了

【図２１】


ファイルアクセス処理229
図21
処理開始
AA:PAからコマンド受信
2101
2102
情報あり？
Yes
No
2103
AA:ディレクトリ情報読み出し、キャッシュへ格納
2104
Read/Write？
No
Yes
2105
Read & キャッシュヒット？
Write
Yes
No
AA:キャッシュ領域確保
2106
2118
当該コマンド処理
2117
2107
連携モード？
Yes
連携モード処理
No
AA:CA/PAへコマンド送信
2108
2109
Write？
Yes
2111
AA:準備完了受信
No
2110
AA:データ受信
2112
AA:データ送信
AA:完了報告受信
2113
No
2114
完了？
Yes
AA:PAへコマンド応答送信
2115
PA:コマンド応答送信
2116
終了

【図２２】

図22


連携モード処理2118
処理開始
AA:PAへ連携モード通知
2201
AA:CA/PAへ連携モードで
コマンド送信
2202
Read？
2203
No
Yes
PA:CA/PAから準備完了受信
2207
PA:CA/PAへデータ送信
2208
PA:CA/PAからデータ受信
2204
AA:完了報告受信
2205
No
完了？
2206
Yes
終了

【書類名】要約書

【要約】

【課題】

NASサーバと当該NASサーバに付与されたネットワークアドレスが設定されるポートとの関係が１：１で一意の決定されるためTCP/IP処理部またはNASサーバの処理性能がネックでレスポンスが低下する。

【解決手段】

ホストがコマンドを送付するアドレスが付与されるI/F部分とNASサーバ部とデバイスの処理を行うキャッシュ制御部と外部ストレージを接続するI/F部分を分離し、相互結合網１７０で接続することで、I/FとNASサーバの対応関係およびI/Fに付与されるアドレスとNASサーバとの対応関係を処理負荷等に応じ任意に変更する。

【選択図】　図１

認定・付加情報

特許出願の番号	特願２００３－２０６１６８
受付番号	５０３０１３０００４２
書類名	特許願
担当官	第七担当上席　　００９６
作成日	平成１５年　８月　７日

＜認定情報・付加情報＞

【提出日】　平成15年 8月 6日

特願２００３－２０６１６８

出　願　人　履　歴　情　報

識別番号　　［０００００５１０８］

１．変更年月日　　１９９０年　８月３１日
［変更理由］　　新規登録
住　所　　東京都千代田区神田駿河台４丁目６番地
氏　名　　株式会社日立製作所